THE PROCEEDINGS OF THE PHYSICAL SOCIETY

VOL. 47, PART 5 *September* 1, 1935 NO. 262

534·323

A CORRECTION TO THE THEORY OF THE RAYLEIGH DISC AS APPLIED TO THE MEASUREMENT OF SOUND-INTENSITY IN WATER

BY A. B. WOOD, D.SC., F.INST.P.

Received March 15, 1935. *Read June* 7, 1935

ABSTRACT. König's theory of the Rayleigh disc assumes the disc to be motionless, apart from its slow rotation. This assumption is unjustifiable in a relatively heavy medium such as water. The disc partakes of the motion of the medium to an extent depending on the density of the medium and the mass and dimensions of the disc. Consequently the fluid-velocity v in König's equation for the torque on the disc must be modified by the introduction of a factor β involving the velocity of the disc relative to the medium. The value of β for ellipsoidal discs is calculated from hydrodynamic theory and determined experimentally. A comparison is made of the results for ellipsoidal discs and for flat cylindrical discs. A method of measuring sound-intensities under water by means of small heavy metal discs is described.

§ 1. THEORY

THE theory of the Rayleigh disc was first developed by W. König[1] in a series of papers entitled "Hydrodynamic Acoustic Investigations". He obtained a general expression for the turning-moment experienced by a stationary ellipsoid of revolution in a vibrating fluid. In the case of an oblate ellipsoid of revolution (a disc-like body) he found that the turning-moment T is given by

$$T = \tfrac{4}{3}\rho a^3 v^2 \sin 2\theta \qquad \ldots\ldots(1),$$

where ρ is the density of the fluid medium, a the radius of disc, v the r.m.s. velocity of the fluid flow and θ the angle between the normal to the disc and the direction of the undisturbed fluid flow. If the disc has appreciable thickness a small multiplying factor $(1 - 0{\cdot}15\ t/a)$, where t is the thickness of the disc, must be applied to the above expression for T. Usually this factor approximates closely to unity.

The streamlines of flow round a disc inclined to the undisturbed stream are illustrated in figure 1. A and B are points of zero velocity and therefore of maximum pressure. The manner in which the torque arises obviously requires no explanation. It will be seen from equation (1) that the couple is a maximum when $\theta = 45°$ and becomes zero when $\theta = 0°$ or 90°, i.e. when the disc is edge-on or broadside-on to the flow. The couple T tending to bring the disc broadside-on to the flow is proportional to v^2, a quantity independent of the direction of flow. Consequently T may be used

to measure v^2 in alternating flow, as in a sound-wave, whence the sound-intensity is calculated.

This method has been in use as a standard method for measuring sound-intensity in air for many years and is still relied upon for the calibration of sub-standards of various kinds. In applying the method to under-water sounds, however, a serious difficulty arises. König's theory, summarized above, assumes the disc to be motionless, apart from its slow rotation, and this assumption is unjustifiable in a relatively heavy medium like water. Allowance must be made for the fact that the disc partakes of the motion of the water, and the quantity v in König's equation (1) must define the velocity of fluid flow *relative* to the disc.* Another point implied, but not definitely stated, in König's theory is that the disc must be rigid, i.e. it must not be set into any form of transverse or radial resonant vibration in the range of frequencies over which it may be used. Otherwise the theory would break down completely.

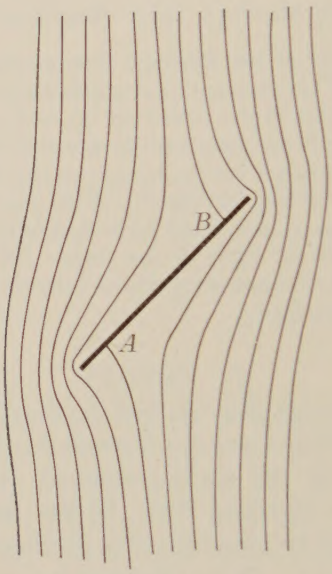


Figure 1. Streamline flow round a thin circular disc inclined at 45°.

v_d, v_w If we represent the r.m.s. velocity of the disc by v_d and of the water by v_w and

β write

$$v_d/v_w = \beta;$$

then the relative velocity v between the water and the disc is given by

$$v = v_w - v_d = v_w(1-\beta) \qquad \ldots\ldots(2).$$

* It is understood from a private communication that Prof. L. V. King, McGill University, Montreal, working independently on this problem, has arrived at the following conclusions as a result of a comprehensive mathematical treatment. "The velocity potential in the neighbourhood of a Rayleigh disc may be completely expressed in terms of cylindrical wave functions. When the circumference of the disc is small compared to the wave-length, the integral equations involved may be approximately solved by the use of Hankel's inversion theorem. The final formula for the torque on the disc differs from Rayleigh's original formula by an inertia factor due to the oscillations of the disc in the incident wave. The final practical formula applicable to the disc as a sound-measuring device shows the existence of an optimum thickness depending on the radius and relative densities of the disc and of the medium in which the sound-waves are propagated."

Equation (1), which gives the couple tending to turn the disc broadside-on, becomes

$$T = \tfrac{4}{3}\rho a^3 v_w^2 (1-\beta)^2 \sin 2\theta \qquad \ldots\ldots(3).$$

We have now to determine the value of β, the ratio of the velocity of the disc to that of the water. If M is the mass of the disc, m that of the water displaced, and m_w the water load, we have

M, m, n

$$\frac{v_d}{v_w} = \frac{m+m_w}{M+m_w} = \beta$$

and

$$(1-\beta) = \frac{M-m}{M+m_w}.$$

It is shown in text-books of hydrodynamics* that the water load on a thin ellipsoidal disc broadside-on to the fluid flow is $\frac{8}{3}a^3\rho_w$. When the disc is inclined at 45° this water load m_w reduces to $\frac{4}{3}a^3\rho_w$ approximately, whilst in the edge-on position the water load on a thin disc approaches zero. When the disc is inclined at 45° we have therefore

$$(1-\beta) = \frac{M-m}{M+m_w} = \frac{M-m}{M+\frac{4}{3}a^3\rho_w} \qquad \ldots\ldots(4).$$

The intensity (erg/cm.^{2} sec.) in the sound-wave in water is given by

$$I = \rho_w C v_w^2 \qquad \ldots\ldots(5),$$

where C is the velocity of sound in water ($1{\cdot}45 \times 10^5$ cm./sec. in fresh water at 15° C.) and v_w is the r.m.s. velocity of the water in the undisturbed sound-wave. Substituting the value of v_w^2 given by equation (3) and assuming $\theta = 45°$ we find that this becomes

C
v_w

$$I = \frac{\frac{3}{4}CT}{a^3(1-\beta)^2} \qquad \ldots\ldots(6).$$

In this expression for the sound-intensity the torque T on the disc in the 45° position can be measured; the quantities a and β are known from the linear dimensions and mass of the disc, and C is the known velocity of sound in the water. It will be shown later how the value of β can be checked experimentally.

A fine suspension wire attached to a torsion head serves to measure the torque T required to maintain the disc in its initial 45° setting against the couple exerted by the sound wave. If α is the angle of twist, $T = k\alpha$, where k is a constant characteristic of the suspension used to support the disc. Small angular deflections near the 45° position may be observed directly without using the torsion head.

α, k

The expression for the sound-intensity given in equation (6) above is of course applicable not only when water is the medium surrounding the disc but also when we are dealing with sounds in air. In the latter case, however, the correction due to the motion of the disc and the mass of air carried with it is small. As an example let us suppose a disc of mica 1 cm. in diameter and 0·002 cm. thick, with M equal to 5×10^{-3} gm., is chosen to measure intensity in air and in water, then the comparative values of the quantities immersed are as shown in table 1.

* Lamb, *Hydrodynamics*, pp. 152, 162 (1895).

Table 1

	Air	Water
M	5×10^{-3} gm.	5×10^{-3} gm.
m	2×10^{-6} gm.	$1{\cdot}6 \times 10^{-3}$ gm.
m_w*	2×10^{-4} gm.	0·16 gm.
β	0·038	0·98
$1/(1-\beta)^2$	1·08	2500

* Load due to the medium (air or water).

The quantity $1/(1-\beta)^2$ is the correction factor required to König's formula for the intensity as measured by the Rayleigh disc. It will be seen therefore that whereas a correction of only 8 per cent is required in the indications of a thin mica disc used in air, the same disc in water would require a multiplying factor of 2500. Such a disc is therefore quite useless in water; a much heavier disc is necessary. For example, compare now the corrections required when the disc employed is a tungsten ellipsoid of mass $M = 6{\cdot}85$ gm. and diameter 2 cm.; these are given in table 2. In this case the factor $1/(1-\beta)^2$ is quite negligible in air and has been reduced to 40 per cent only in water.

Table 2

	Air	Water
M	6·85 gm.	6·85 gm.
m	$4{\cdot}7 \times 10^{-3}$ gm.	0·36 gm.
m_w	$1{\cdot}1 \times 10^{-3}$ gm.	0·85 gm.
β	9×10^{-4}	0·158
$1/(1-\beta)^2$	1·002	1·4

These two examples serve to show the necessity for using as heavy a disc as possible in water. In air this necessity is far less urgent, but it appears that the correction factor may become important if very light mica discs are used.

I The acoustic power-output of any oscillator may be determined when the axial intensity I and the beam-shape are known. In certain cases, as for example a high-frequency circular oscillator of piston type, the beam-shape conforms to a known theoretical form[2, 3] whence the sound-output is easily calculated in terms of the measured sound-intensity on the axis. The total acoustic energy-output per E, S second E of such an oscillator of area S vibrating with r.m.s. velocity v over its whole surface is given by

$$E = \rho c S v^2 \qquad \text{.....(7),}$$

λ provided the diameter of the oscillator is considerably greater than a wave-length λ. R_c Such a source emits an approximately parallel beam of radiation to a distance R_c, where

$$R_c = m^2 \lambda$$

m, r and $m = r/\lambda$, r being the radius of the circular oscillator. For values of R appreciably larger than R_c the axial intensity I is given by

$$I = \rho C S v^2 \pi m^2 / R^2 \qquad \text{.....(8),}$$

whence from (7) and (8)

$$E = \frac{\lambda^2 R^2}{S} I \qquad \text{.....(9).}$$

The value of I is determined by means of the disc at a distance R from the oscillator by means of equation (6) above, whence the power-output E of the sound-source is readily calculated from equation (9).

If the beam-shape of the oscillator does not conform to a standard type it may be necessary to perform a three-dimensional integration to obtain the power-output in terms of the axial intensity.

For a non-directional source whose dimensions are small compared with the wave-length λ, the power-output E is given by

$$E = 4\pi R^2 I \qquad \text{.....(10).}$$

§ 2. EXPERIMENTAL MEASUREMENTS

Determination of the water load m_w. The value of the water load on the disc may be simply calculated; in the 45° position of the disc, as we have seen, $m_w = \frac{4}{3}\rho_w a^3$. It is nevertheless of interest and importance to provide an experimental check on the theoretical value. Equation (4) above may be re-written in the form

$$m_w = \frac{M-m}{1-\beta} - M \qquad \text{.....(11).}$$

For sounds of constant intensity the torque on the disc is proportional to $a^3(1-\beta)^2$ as shown by equation (6). For discs of the same radius a, therefore, the angular deflection δ of the disc is proportional to $(1-\beta)^2$ or $\sqrt{\delta}$ is proportional to $(1-\beta)$. Equation (11) may therefore be written

$$m_w = K\left(\frac{M-m}{\sqrt{\delta}}\right) - M \qquad \text{.....(12),}$$

where K is a constant.

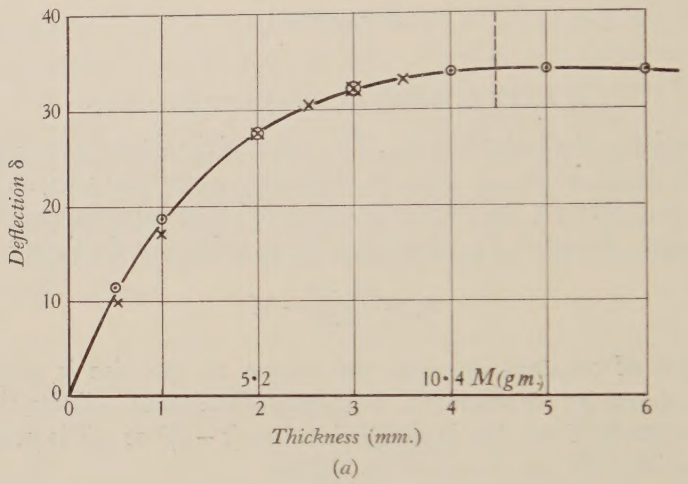
Consequently if deflections δ are measured for various values of M, the graph of $(M-m)/\sqrt{\delta}$ and M will be a straight line cutting the axis of M at $-m_w$. For discs of the same radius, the value of M may be varied (i) by varying the thickness of discs of the same material, and (ii) by using materials of different densities.

The results of such measurements are shown in figures 2 and 3. In figure 2, cylindrical brass discs of radius 1 cm. and of varying thickness were used. In figure 2*a* the observed values of angular deflection δ are plotted as a function of thickness, which is proportional to M, whilst in figure 2*b* the values of $(M-m)/\sqrt{\delta}$ are plotted against M. A similar pair of curves is shown in figures 3*a* and 3*b*. In this case cylindrical discs of different density but of the same dimensions (radius 1 cm., thickness 0·2 cm.) were used. It will be seen that both figures 2*b* and 3*b* indicate a linear relationship between M and $(M-m)/\sqrt{\delta}$ and yield a value of 1·3 gm. for m_w. Similar observations plotted in figure 5, to which reference will be made later, give a value 1·1 gm. These experimental values agree tolerably well with the theoretical estimate of $\frac{4}{3}$ gm. for ellipsoidal discs of radius 1 cm.

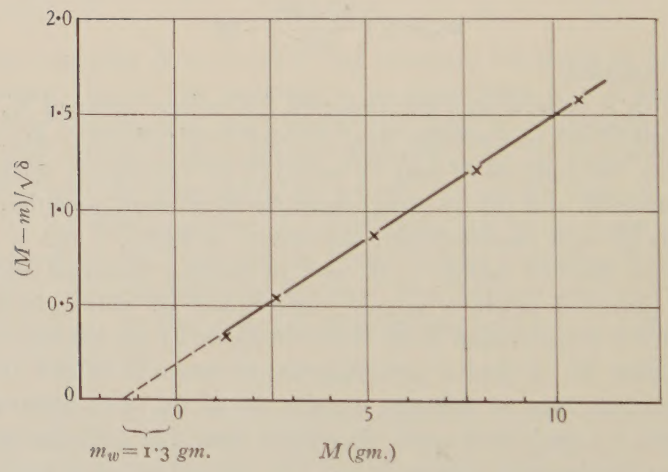
A comparison has also been made of discs of different radii but of the same material and thickness, figure 4. The calculated and observed values are compared in table 3.

Table 3

Diameter $2a$ of disc (cm.)	$1-\beta$	$a^3(1-\beta)^2$	Relative deflections (degrees)	
			Calculated	Observed
1·5	0·748	0·235	1	1
2·0	0·707	0·500	2·13	2·3
2·5	0·675	0·891	3·8	3·8



(a)



(b)

Figure 2. Cylindrical brass discs of radius 1 cm. and of varying thickness.

In the calculation of the fourth column of this table the theoretical value $\frac{4}{3}\rho_w a^3$ has been assumed for the water load on the discs. The agreement between observation and calculation gives support to this value and to the theoretical prediction that the deflection is proportional to the cube of the radius of the disc.

In the above experiments the discs used were all cylindrical and flat. The theoretical deductions of König relate, however, to ellipsoidal discs, and it is of some importance to examine what difference, if any, is introduced by the use of cylindrical

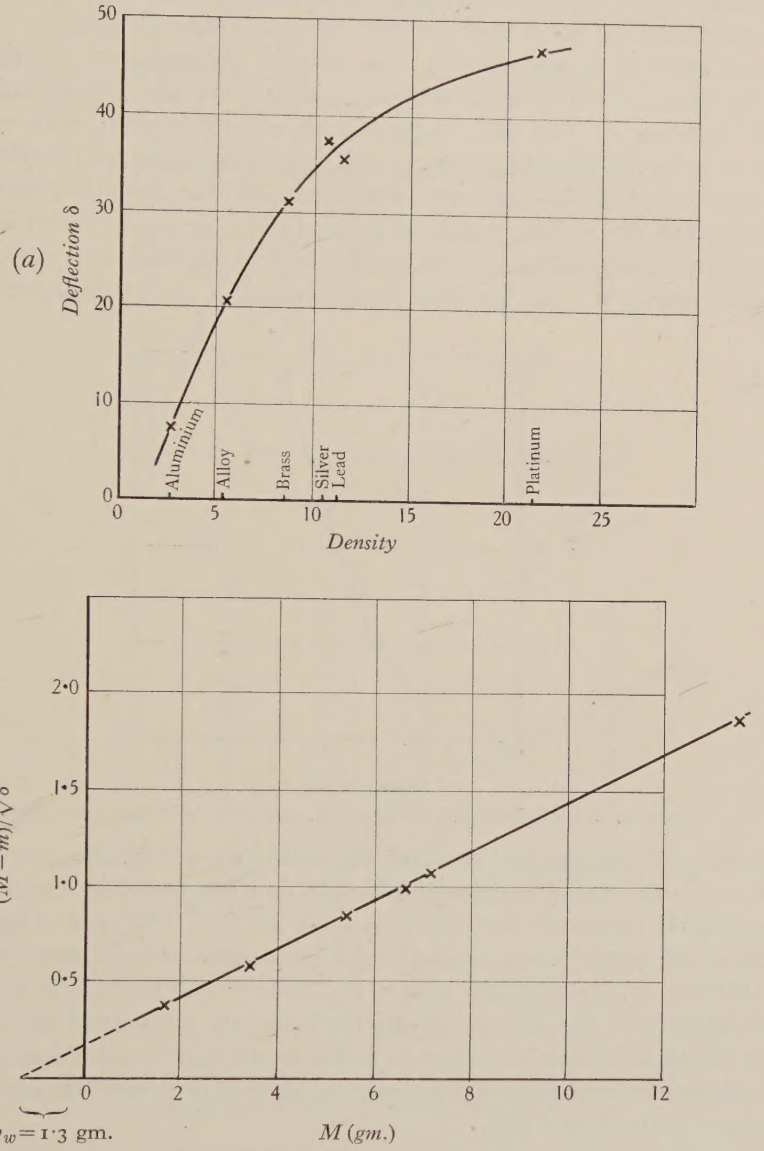


Figure 3. Cylindrical discs of different materials, radius 1 cm., thickness 2 mm.

discs. A comparison was therefore made between a series of cylindrical brass discs of radius 1 cm. and of varying thickness, and a similar series of ellipsoidal discs. The results, plotted in figure 5, indicate that the difference if any is small, the slope of the line and the m_w intercept being practically the same in the two cases. Differences

arise when the thickness of the discs exceeds 4 mm. but this would be expected on theoretical grounds.

In accordance with theory, the results shown in figures 2, 3 and 5 clearly indicate the increased sensitivity obtainable by increasing the weight of the disc. The correction factor due to the water load can thus be made relatively small.

Measurement of sound-intensity. The general arrangement of apparatus for measuring sound-intensity under water by means of the disc is illustrated diagrammatically in figure 6. The disc, suspended on a short (10-cm.) wire of phosphor bronze, 0·028 cm. in diameter, is supported from a long stiff rod carrying a torsion head. In the first experiments a spot of light was reflected from the polished surface of the disc, but the surface polish of metal discs is as a rule very poor and this method had to be abandoned. Considerable optical improvement was obtained by using a thin (0·010-in.) stainless steel concave mirror 1 cm. in diameter mounted at

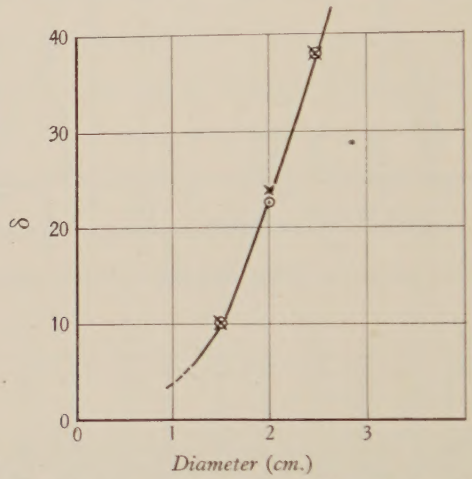


Figure 4. Cylindrical brass discs 2 mm. thick and of varying diameter.

an angle of 45° to the disc and arranged edge-on to the sound. The torque on such a mirror due to the sound is theoretically zero and was found by experiment to be quite negligible compared with the torque on the disc. The spot of light reflected from the small mirror was arranged to fall on a white xylonite scale where it was easily visible to an eye vertically above. The scale served not only to indicate small angular deflections, but also to indicate the exact 45° position of the disc. In this position the small mirror was edge-on to the sound and broadside-on to the light, a very convenient arrangement, the beam of light to and from the mirror being at right angles to the sound-path.

The torsion head controlling the position of disc could be operated by means of a pulley and cord situated conveniently for the observer near the under-water lamp and scale. An inclined mirror over the torsion head was used to reflect towards the observer an image of the pointer and angular scale of the suspension. To reduce the effects of direct flow of water past the disc, due to possible disturbances in the tank, a thin (0·010-in.) clear celluloid cylinder about 6 in. in diameter was arranged as

shown in figure 6. This flabby celluloid screen, whilst sufficient to stop slowly-varying water currents, permitted the sound to pass freely, observations with and without the cylinder showing no appreciable difference. Provision was also made for transport of the disc without risk of breaking the suspension—a small metal cone

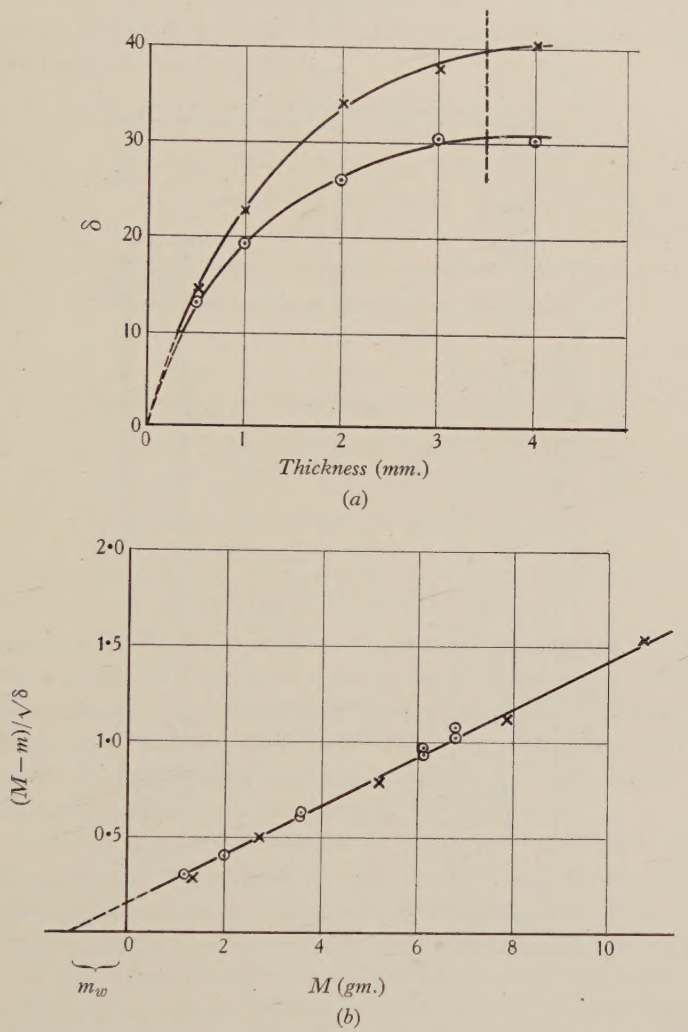


Figure 5. Comparison of cylindrical and ellipsoidal brass discs of radius 1 cm. Ellipsoidal discs ⊙ ⊙; cylindrical discs × ×.

at the lower end of a rod served to raise the disc when required and consequently to remove the tension from the fine suspension wire. Experiments showed that increased damping of the suspended system could be obtained, if required, by means of a thin mica vane mounted symmetrically on the disc-suspension rod and arranged either edge-on or broadside-on to the sound. In these positions the torque on the damping vane due to the sound is zero. Generally, however, the damping due to the

disc and mirror was found sufficient without the addition of the mica damping-vane. The torsion constant of the suspension was determined experimentally by observing the period of torsional oscillation in air when a disc of known mass and radius was suspended from a point on its edge. The torque on the disc is given by

$$T = k\alpha,$$

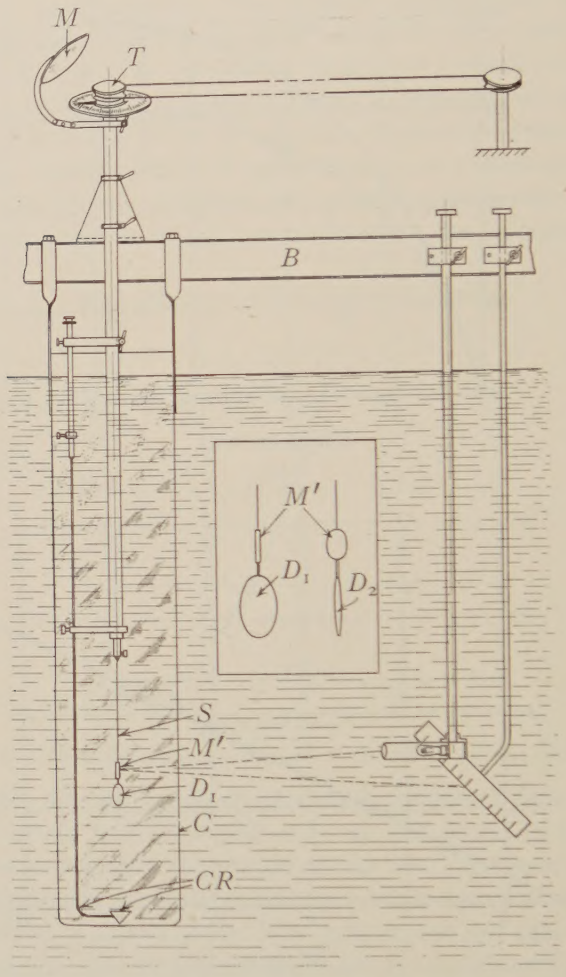


Figure 6. General arrangement. M, M', mirrors; T, torsion head and scale; B, bridge; D_1, disc at 45° to sound; D_2, disc edge-on to sound; CR, clamping rod and cup; C, thin celluloid cylinder; S, suspension wire.

k, τ where k is a constant equal to $\pi^2 Ma^2/\tau^2$, M the mass in grams of the cylindrical disc of radius a, τ the period in seconds of torsional oscillation and α the angle of twist measured in radians.

In using this apparatus to measure sound-intensity in the tank it was soon discovered that small changes of frequency in the source of sound produced appreciable changes in deflection of the disc. This effect was traced to stationary waves formed

by reflection from the walls of the tank. The difficulty was surmounted by superposing a small frequency-modulation on the transmitted sound. In this connection it is important also to note that the disc has directional properties, which make it more sensitive to sound arising in some directions than in others. The couple T turning the disc is proportional to $\sin 2\theta$, where θ is the angle between the direction of the sound and the normal to the disc. The directional characteristic shown in figure 7 shows maxima at $45° \pm n\pi/2$. θ

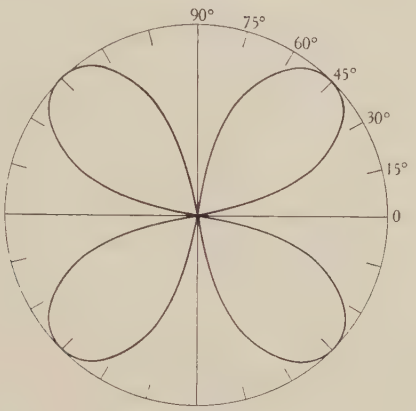


Figure 7. Directional characteristic of disc.

§ 3. SIZE AND FREQUENCY OF TRANSVERSE VIBRATION OF THE DISCS

In the theory of the Rayleigh disc it is stipulated that the circumference of the disc should be considerably less than a wave-length of the sound of which the intensity is to be measured. In water this condition is easily fulfilled since the velocity of sound is high, 1450 m./sec., a sound of frequency 1000 c./sec. having a wave-length of 145 cm. The small discs 1 cm. in radius used in the present experiments are in this respect in accord with the theoretical conditions.

Another factor of importance, which is generally overlooked, relates to the possibility of transverse vibration of the disc in its fundamental mode, with two nodal diameters, or at its first overtone with one nodal circle. When a disc is inclined at 45° to the direction of the sound-wave a pressure-difference exists between the forward and rear edges of the disc, and if the frequency of sound happens to coincide with one of the natural resonant frequencies of the free disc a vibration may be set up which renders the indications valueless. The theoretical frequencies of free circular discs in air are given by Rayleigh* as follows. For a disc vibrating with two nodal diameters (the lowest possible natural frequency of the disc)

$$N_x = 0{\cdot}26\, t\, c/a^2, \qquad N_x$$

where t is thickness and a the radius of the (cylindrical) disc and c the velocity of sound in the material of the disc. For a disc vibrating with one nodal circle

$$N_0 = 0{\cdot}41\, t\, c/a^2. \qquad N_0$$

* Vol. 1, chap. x. See also Lamb, *Sound*, p. 153.

These expressions refer to flat cylindrical discs. The frequencies of ellipsoidal discs of the same radius and thickness at the centre will of course be considerably higher, owing to removal of mass from the outer edge without much change of stiffness.

In an accompanying paper a description is given of the experimental determination of the frequencies of small circular discs, and it is shown that the experimental values are in reasonably good agreement with the theory. For the present therefore we may use the values of frequency calculated from the above expressions.

The lowest frequency of a cylindrical platinum disc, 2 cm. in diameter and 2 mm. thick, was found to be about 14,000 c. sec., whilst that of an ellipsoidal tungsten disc 2 cm. in diameter and 2 mm. thick at the centre was higher than 20,000 c. sec. The frequencies in water will be slightly lower than these values. Such discs are therefore quite safe to use over a wide range of frequencies extending into the supersonic region.

In this connection it is of interest to calculate the frequency of a small mica disc such as is commonly used to measure sound-intensity in air. If the disc has a radius of 0·4 cm. and a thickness of 0·0025 cm., it could vibrate with two nodal diameters at a frequency N_x where

$$N_x = 1300 \text{ c./sec.}$$

and with one nodal circle at a frequency N_0 where

$$N_0 = 2050 \text{ c./sec.}$$

In the calculation of these frequencies the velocity of sound in sheet mica was taken as $3{\cdot}2 \times 10^5$ cm./sec., a value found experimentally by a method described by the author and Dr F. D. Smith[4]. It will be seen that the small mica disc has two natural modes of vibration within the aural frequency range.

Data relating to two discs for use under water. The data incorporated in table 4 relate to a cylindrical platinum disc and an ellipsoidal tungsten disc which have been used in the measurement of sound-intensity under water.

Table 4

	Platinum disc, cylindrical	Tungsten disc, ellipsoidal
Radius	1·00 cm.	0·863 cm.
Thickness (at centre)	0·20 cm.	0·216 cm.
Frequency (2 nodal diameters)	14 kc./sec.	20 kc./sec.
Mass M	13·5 gm.	6·85 gm.
Mass of water displaced m	0·63 gm.	0·36 gm.
Calculated water load m_w	1·33 gm.	0·85 gm.
$1-\beta$	0·87	0·842
$1/(1-\beta)^2$	1·32	1·41
v_w^2	$0{\cdot}99\ k\alpha \times \pi/180$	$1{\cdot}64\ k\alpha \times \pi/180$
I (erg/sec.-cm.^{2})	$1{\cdot}44 \times 10^5 k\alpha \times \pi/180$	$2{\cdot}39 \times 10^5 k\alpha \times \pi/180$
I (erg/sec.-cm.^{2}) when $k = 0{\cdot}225$*	565 α†	940 α†

* With an annealed phosphor-bronze wire suspension 28 μ in diameter and 10 cm. long, k is 0·225.

† α is in degrees.

§4. MEASUREMENT OF SOUND-INTENSITY IN WATER. SOUND-OUTPUT OF SOURCES

An example of the type of measurement made by means of the discs described above is illustrated in figure 8. In this case the sound-intensity at a distance of 1 m.

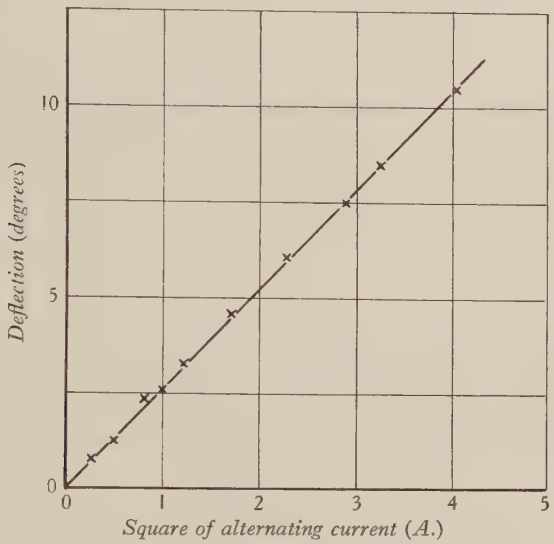


Figure 8.

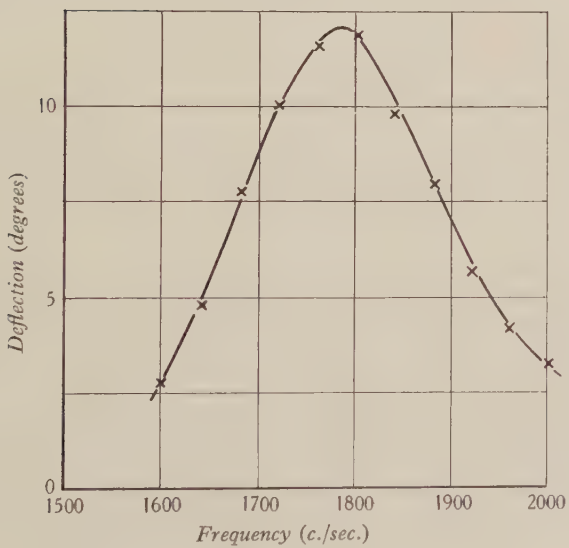


Figure 9.

from the source of sound is plotted as a function of the square of the value of the alternating current used to excite the source into vibration, i.e. as a function of power-input. It will be seen that the sound-output is directly proportional to the

power-input. With an input current of 2 A., it is found that the sound-intensity at 1 m. is 6800 erg/cm² sec. Assuming that the sound-intensity is uniform in all directions, then the acoustic power-output of the source is given by

$$E = 4\pi \times 100^2 \times 6800 \text{ erg/cm}^2 \text{ sec.}$$
$$= 85 \times 10^7 \text{ erg/cm}^2 \text{ sec.}$$
$$= 85 \text{ W.}$$

In another case the disc has been used to determine the frequency characteristic of a sound source under water. This type of measurement is illustrated in figure 9. The intensity of the sound emitted is shown throughout a range of frequencies embracing the resonance frequency of the sound-source.

§ 5. AN ALTERNATIVE TO THE RAYLEIGH DISC—THE DOUBLE SPHERE

W. König[1] not only derived an expression for the torque on a thin ellipsoidal disc, but also for the torque on an elongated prolate ellipsoid, approximating to a cylindrical rod of diameter small compared with its length supported horizontally from the mid point. Such a rod, adjusted initially with its axis at 45° to the direction of the sound-wave, is subjected to a couple tending to rotate it into the broadside-on position.

König[5] also calculated the forces operating on two spheres in an alternating fluid flow and Rayleigh* has pointed out that in the case where the two spheres are rigidly connected together the forces reduce to a couple T where

$$T = 3\pi\rho \frac{a_1^3 a_2^3}{r^3} v^2 \sin 2\theta,$$

θ — θ being the angle which the axis of the combination makes with the sound direction,
a_1, a_2, r, v — a_1 and a_2 the radii of the spheres, r the distance apart of their centres, v the r.m.s.
ρ — velocity of the fluid flow and ρ the density of the fluid. When $a_1 = a_2 = a$ and $r = ma$,

$$T = \frac{3\pi}{m^3} \rho a^3 v^2 \sin 2\theta.$$

It will be seen that this expression is of the same form as that obtained by König for the Rayleigh disc. The couple is proportional to v^2 (which in turn is proportional to the sound-intensity) and is a maximum when the axis of the combination is inclined at 45° to the direction of the sound. The sensitivity of the arrangement varies directly as a^3 and inversely as m^3. The theory requires, however, that the spheres should not be too close together, i.e. m must be considerably greater than 2. In comparing this arrangement with the Rayleigh disc the factor $3\pi/m^3$ in the expression for the torque must be considered in relation to the corresponding quantity $\frac{8}{3}$ for the disc. For the double-sphere combination, m can never be less than 2 when the spheres are in contact and the constant $3\pi/m^3 = 1{\cdot}18$. If $m = 4$, a more likely arrangement, $3\pi/m^3 = 0{\cdot}147$, which is only $\frac{1}{18}$ of the constant for a disc of the same radius.

* *Sound*, II, 47.

As a matter of general interest, a few preliminary experiments have been made with two pith balls 0·25 cm. in radius mounted with their centres 1·0 cm. apart. In this case $a = 0{\cdot}25$ and $m = 4$. The combination was suspended by a thin silk fibre in a tunable resonance tube, the double-sphere arrangement replacing the usual Rayleigh disc. The experimental results were found to agree, within the limits of the rough comparison, with the theoretical value for ratio of sensitivities of the double sphere and the disc.

Although the double-sphere arrangement is always less sensitive and perhaps less convenient in air than the disc, it is just possible that it may have compensating advantages in its application to the measurement of sound-intensity under water. As yet, however, there is insufficient information regarding the water load and relative velocity-amplitude of such an arrangement to warrant its use as a standard of intensity-measurement.

§ 6. ACKNOWLEDGMENT

In conclusion the author desires to thank the Admiralty for permission to publish this paper.

REFERENCES

(1) König, W. *Ann. Phys.*, Lpz., **43**, 43 (1891).
(2) Crandall, L. B. *Vibrating Systems and Sound*, p. 137.
(3) Boyle, R. W. *Trans. roy. Soc. Can.* **20**, 233 (1926).
(4) Wood, A. B. and Smith, F. D. *Proc. phys. Soc.* **47**, 149 (1935).
(5) König, W. *Ann. Phys.*, Lpz., **42**, 549 (1891).

534.12

AN EXPERIMENTAL DETERMINATION OF THE FREQUENCIES OF FREE CIRCULAR PLATES

By A. B. WOOD, D.Sc., F.Inst.P.

Received March 15, 1935. *Read June* 7, 1935

ABSTRACT. The paper deals with an experimental determination of the frequency of vibration of circular plates of appreciable thickness. Two sets of observations made (i) with steel discs varying in diameter from 6 to 16 cm. and in thickness from 0·07 to 2·8 cm., and (ii) with discs of various materials all 2 cm. in diameter but varying in thickness, are described. Frequencies were measured at which the discs vibrated with (*a*) two nodal diameters, (*b*) one nodal circle. The experimental results are compared with theoretical values. The method may be applied to determine the velocity of sound in materials available only in small quantities.

§ 1. INTRODUCTION

THE complete theory of the vibration of a free circular plate was developed by Kirchhoff in 1850. Before this (in 1829) Poisson had investigated theoretically some of the symmetrical modes of vibration and calculated the diameters of the nodal circles. Rayleigh[1] gives the general theory and quotes Kirchhoff's observations on the low frequencies of thin plates, and Lamb[2] summarizes the various theoretical results.

N It is stated that for a plate of known lateral dimensions the frequency N is given by

$$N = 0{\cdot}0462\, ctm^2 \qquad \ldots\ldots(1),$$

c, t where c is the velocity of sound appropriate to the material, t the thickness of the
m plate and m a constant for a plate of given dimensions. The lowest natural frequency of vibration is that in which the free circular plate has two nodal diameters, viz.

$$N = 0{\cdot}261\, ct/a^2 \qquad \ldots\ldots(2),$$

a where a is the radius of the plate. According to Poisson the mode of vibration having one circular node has a frequency N given by equation (1) in which

$$m^2 = 8{\cdot}89/a^2,$$

whence

$$N = 0{\cdot}410\, ct/a^2 \qquad \ldots\ldots(3).$$

r The radius r of the nodal circle is given by

$$r = 0{\cdot}678\, a \qquad \ldots\ldots(4).$$

It will be seen from relations (2) and (3) that the frequency varies directly as the velocity of sound and the thickness of the plate and inversely as the square of the radius.

So far as I am aware no information has been published relative to the vibrations of free circular plates of appreciable thickness. Such plates have high natural

frequencies of transverse vibration which may lie outside the audible range. The present paper is mainly concerned with such plates. The experimental observations are conveniently divided into two series which deal with (i) circular steel plates ranging in diameter from 6 to 16 cm. and in thickness from 0·07 to 2·8 cm., and (ii) small circular plates 2 cm. in diameter and made of various materials.

§ 2. CIRCULAR STEEL PLATES

In this series of experiments the resonant frequency of the steel discs was determined by exciting them into vibration by means of a pot electro-magnet. The centre pole of this magnet, a bundle of soft iron wires, carried at its upper end the alternating-current winding, whilst a direct-current winding was provided in the usual manner. The steel discs rested in a horizontal position on a thin soft pad interposed between the disc and the face of the electromagnet. Sand was sprinkled over the surface of the disc and the frequency of the alternating current varied until the nodal diameters or nodal disc appeared. The tuning was found in all cases to be very sharp, so that the resonant frequency, determined by means of a calibrated wave-meter (or a monochord in the case of the lower frequencies), was accurately known. Observations in this series were confined almost entirely to the frequency at which the nodal circle appeared; in two cases only were the frequencies corresponding to two nodal diameters recorded. The more important results are included in table 1.

Table 1. Mild steel discs vibrating with one nodal circle

Number	Diameter $2a$ (cm.)	Thickness t (cm.)	t/d	Frequency (c./sec.)	Diameter $2r$ of nodal circle (cm.)	r/a
1	6·3	1·00	0·159	19,950	4·23	0·670
	6·3	1·00	0·159	(12,300—2 nodal diameters)		—
2	6·82	0·95	0·139	16,500	4·60	0·674
3*a*	7·00	1·45	0·207	21,800	4·65	0·665
	7·00	1·45	0·207	(13,750—2 nodal diameters)		—
3*b*	7·00	1·20	0·172	19,000	4·65	0·665
3*c*	7·00	1·00	0·143	16,400	4·65	0·665
3*d*	7·00	0·95	0·135	15,800	4·65	0·665
4*a*	8·23	1·07	0·130	12,900	—	—
4*b*	8·23	1·02	0·124	12,650	5·55	0·673
5	8·80	0·125	0·0143	1,480	5·9	0·670
6	8·90	1·50	0·169	14,800	5·95	0·670
7	8·90	0·0735	0·0083	845	6·0	0·675
8*a*	8·90	1·80	0·204	16,900	5·95	0·670
8*b*	8·90	1·50	0·169	14,800	5·95	0·670
8*c*	8·90	1·00	0·1125	10,650	6·0	0·675
8*d*	8·90	0·60	0·067	6,700	6·0	0·675
9*a*	15·5	2·77	0·178	8,750	10·4	0·670
9*b*	15·5	1·0	0·065	3,700	—	—

It will be seen in all cases that the ratio r/a of the radius of the nodal circle to the radius of the disc shows very little variation with the thickness and diameter of the discs. The extreme values are 0·665 and 0·675 whilst the theoretical value obtained by Poisson for a thin plate is 0·678. The experimental result is all the more remark-

able when it is considered that the ratio of thickness to diameter of the discs in some cases reaches a value as high as 1 : 5.

With regard to the frequency, however, the results are not quite so simple. Figure 1 (*a*) shows the relation between thickness and frequency for discs of different

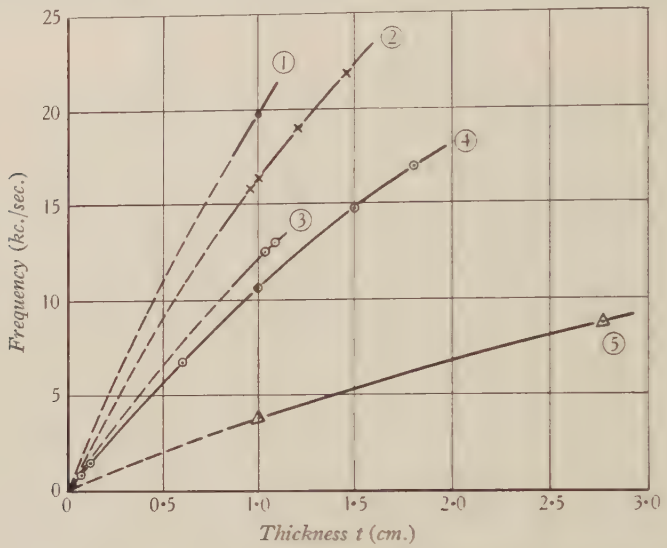


Figure 1 (*a*). Diameters ① 6·3 cm.; ② 7·0 cm.; ③ 8·23 cm.; ④ 8·90 cm.; ⑤ 15·5 cm.

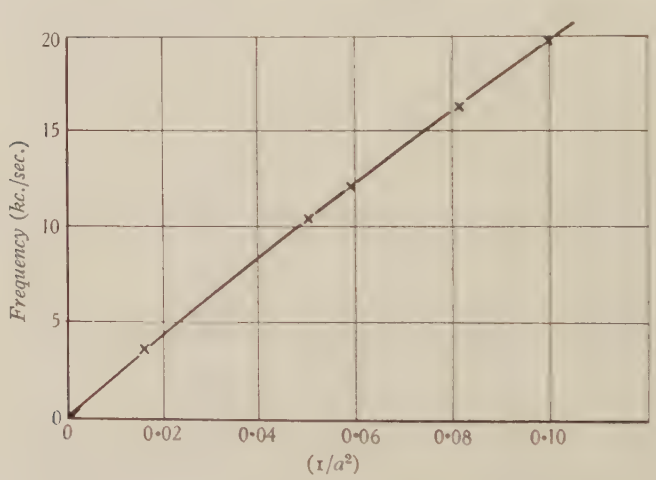


Figure 1 (*b*).

diameters. The frequency increases at a rate which is slightly less than proportional to thickness. Similarly in figure 1 (*b*) the relation between frequency and $1/a^2$ is not quite linear. These results indicate a small but definite departure from the theoretical relationship indicated by equation (3) when the plates can no longer be regarded as thin.

The observations for the thinnest plates, nos. 5 and 7 in table 1, lead to a value of $5{\cdot}6 \times 10^5$ cm./sec. for the velocity of sound in thin sheet steel, if the theoretical value 0·410 for the numerical constant in equation (3) is correct. On this basis the values of this constant K have been calculated for the thicker discs. The results are plotted in figure 2. The points all lie fairly well on a smooth curve in spite of the fact that the thickness and diameter of the discs are varied, independently, over a wide range. It will be seen that the value of K decreases fairly rapidly as the thickness of the disc increases, relatively to the diameter.

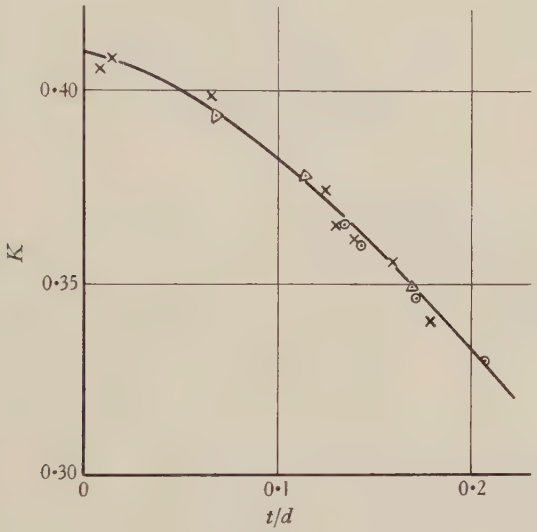


Figure 2.

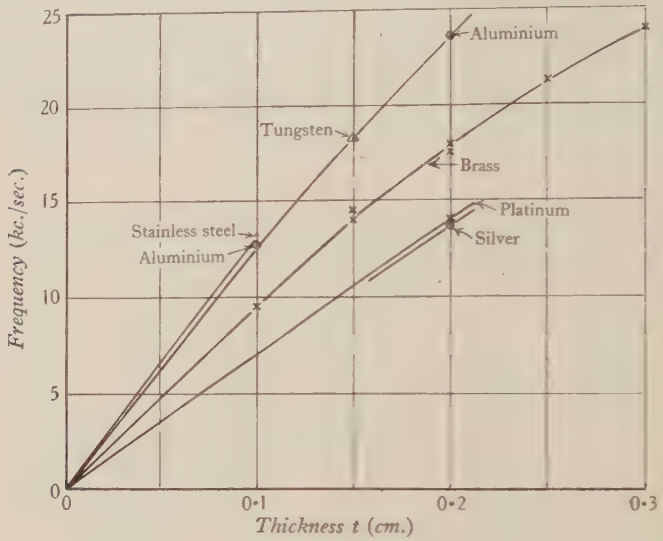
§ 3. OBSERVATIONS WITH SMALL CIRCULAR PLATES OF VARIOUS MATERIALS

In connection with the experiments described in the accompanying paper relating to measurement of sound-intensity in water by means of a Rayleigh disc, it was important to determine the natural frequencies of vibration of free discs 2 cm. in diameter. These discs were made of materials varying in density from 2·65 (aluminium) to 21·5 (platinum). The only available information relative to the frequency of vibration of such discs was of a purely theoretical nature, e.g. equations (2) and (3) above, and involved a knowledge of the velocity of sound in the material. Tabulated information bearing on this point was scanty and often unreliable. It was considered desirable therefore to determine the frequencies of such discs experimentally.

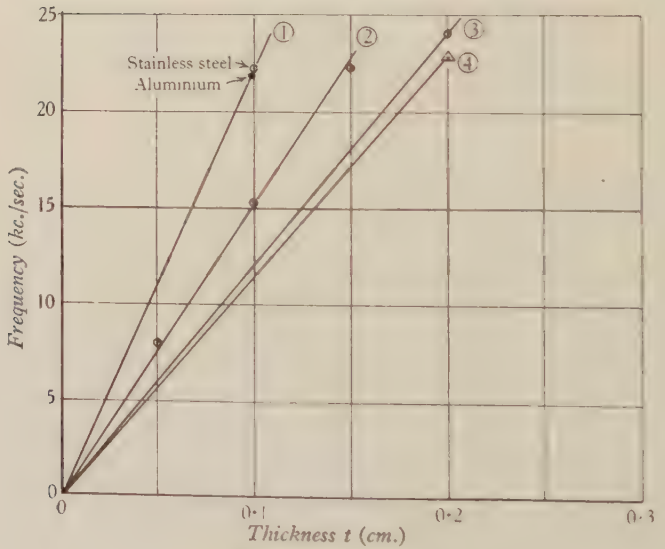
The method described in the previous section of the paper with reference to steel discs was inapplicable to this case, for most of the discs were non-magnetic. A simple and reliable method was devised in which the source of vibration of variable frequency was a small magnetostriction oscillator.* This consisted of a

* As used in echo depth recording. See British patent specification No. 375,375.

laminated nickel block supplied with a polarizing-winding carrying direct current, and an alternating-current winding connected to a valve oscillator having a wide range of frequency-control.



(*a*) Two nodal diameters.



(*b*) One nodal circle.

Figure 3. Frequencies of metal discs (radius $a = 1$ cm.). ① aluminium and stainless steel; ② brass; ③ platinum; ④ silver.

The small disc under examination was placed on the horizontal vibrating surface of the magnetostriction oscillator and sprinkled with fine sand. The frequency of excitation was then adjusted until the required sand figure, two nodal diameters or

one nodal circle, appeared. The frequency was then observed in the usual manner by means of a calibrated wave-meter. The results of such observations are shown plotted in figure 3 (*a*) and (*b*) where the frequencies, ranging from 8 kc./sec. to 23 kc./sec., are plotted as a function of thickness, (*a*) relating to two nodal diameters and (*b*) to one nodal circle. All the discs were 2 cm. in diameter.

As in the case of the large steel discs, the frequency increases at a rate which is slightly less than proportional to the thickness. The radius of the nodal circle was found to approximate closely to $0{\cdot}67a$. The dimensions and frequencies of each disc being known, the velocity of sound may be calculated from relations (2) and (3) above, the constants 0·261 and 0·410 being assumed. Velocities determined in this way and included in table 2 are therefore subject to a correction depending on the ratio of thickness to diameter as indicated in figure 2.

Table 2

Material	Velocity of sound (cm./sec.)	
	(*a*) from observations of two nodal diameters	(*b*) from observations of one nodal circle
Brass	$3{\cdot}6 \times 10^5$	$3{\cdot}7 \times 10^5$
Aluminium	4·9 ,,	5·4 ,,
Stainless steel	5·1 ,,	5·5 ,,
Silver	2·6 ,,	2·8 ,,
Tungsten	4·6 ,,	—
Platinum	2·7 ,,	2·9 ,,

There appears to be a tendency for the nodal-circle observations to yield somewhat higher values of velocity than those derived from the nodal diameters. Considering the quantity of material available in discs 2 cm. in diameter and 1 or 2 mm. in thickness, these experimental values for the velocity of sound are, however, in remarkably good agreement with those obtained in larger-scale determinations. The method shows promise therefore as a means of determining the velocity of sound in materials available only in small quantities, e.g. in the form of a small thin disc or button.

§ 4. ACKNOWLEDGMENT

In conclusion the author desires to thank the Admiralty for permission to publish this paper.

REFERENCES

(1) Rayleigh. *Sound*, I, Chap. X.
(2) Lamb. *Dynamical Theory of Sound*, p. 152.

53.088.2

THE STATISTICAL THEORY OF ERRORS

By NORMAN CAMPBELL, Sc.D., F.Inst.P.

Received February 14, 1935. *Read in title June* 7, 1935

ABSTRACT. It is maintained that the adjustment of observations is *not* a statistical problem. In § 1 the problem is stated. In § 2 it is argued that there is no proof that the conditions necessary for a statistical theory of errors exist. In § 3 it is argued that the process of adjustment adopted by experimenters can be justified apart from statistical considerations. In § 4 the conclusions thus attained are reviewed. The actual argument cannot be summarized.

§ 1. INTRODUCTION

Deming and Birge[1] have recently discussed at great length the application of modern statistical methods to the theory of errors. They are probably not so sanguine as to believe that their work will have much effect on the practice of experimental physicists. It is as certain as anything can be that the great majority of us will fail to use the weapons they offer, and will continue to employ only the very simple methods of adjusting observations that are familiar to us; it is equally certain that this great majority will include almost all of the greatest masters of the experimental art.

And yet, if Deming and Birge's opening statement is true, we and our masters must be wrong. If "problems that arise in drawing conclusions from observations are essentially statistical", then it cannot be right to use for the solution of such problems any but the most perfect methods known to modern statisticians. If we are to justify ourselves, we must attack the statistical theory of errors much more fundamentally than heretofore; we must cease to object to particular statistical methods and assumptions, and boldly deny from the outset Deming and Birge's fundamental statement. We must deny that the adjustment of observations is a statistical problem at all.

A frivolous analogy may be useful. A golfer sees some men hitting a ball about with bent sticks. Noticing that they are urging the ball in a prescribed direction, and that their sticks are not unlike those shown in 18th-century golf pictures, he concludes that they are trying to play golf. He upbraids them for their ignorance of the modern art, and expounds to them the virtues of rubber-cored balls, steel-shafted drivers and a panoply of irons. They have no defence if they merely say that modern golf is over-elaborate, or that their implements are good enough for the very crude form of the game they are playing; to silence him they must reply that they are not playing golf at all, but hockey; they must point out that their game, though presenting certain superficial resemblances to golf, is so fundamentally different that his instructions are not wrong, but irrelevant.

§ 2. STATISTICS

Statistics, or at least all the statistical methods of Deming and Birge's paper, are directed to obtaining information about a collection* from a sample drawn from it. The problem is trivial unless the members of the collection are known to differ amongst themselves, so that members of the sample may also differ. In the adjustment of observations we are concerned with differing measurements; this is the fact that has led statisticians to believe that such adjustment is a statistical problem. But this belief is not justified unless (*a*) the differing observations are a sample of a collection, the sample and the collection being of the kind with which statistics is concerned; and (*b*) complete knowledge of the collection would solve or help to solve the problem of adjustment. In this section we shall consider (*a*), leaving (*b*) till later.

If measurements M to be adjusted are a sample of a collection, this collection can only be one of three things. It must be either N_1, N_2 or N_3, where $N_1 = M +$ all other similar measurements that are known; $N_2 = N_1 +$ all the similar measurements that will be known; and $N_3 = N_2 +$ all the similar measurements that might be known. If the collection is N_1, the problem can be, and ought to be, removed at once from the realm of statistics by discussing not the sample M but the whole collection N_1. If it is N_2 it can be, and ought to be, similarly removed by proceeding at once by making such a complete series of measurements that nobody will ever want to add to it. Accordingly the problem is essentially statistical only if the collection is N_3. This is apparently Deming and Birge's view. It does not seem to me at all sure that N_3 means anything at all; I cannot clearly conceive of a measurement that might be or might have been made and yet never will be or has been made. But the discussion may proceed on the assumption that the collection is N_3.

M
N_1
N_2
N_3

We have next to inquire what are the samples and collections to which statistical theory is applicable. One answer, which might have been given at the end of the 18th century, is that it is applicable to all samples of all collections. Nobody surely will give that answer to-day; everyone recognizes that it is absurd to treat statistically the answers of a jury in a criminal trial. The reasons why that answer is false have been given very fully in our own day by Keynes. But beyond that there is no agreement. Statistical methods depend on the conception of a measurable probability†, and there is as yet no agreement as to what propositions have probability and how they come to have it. Progress has been made, for example by Mises and by Jeffreys, in determining what mathematical axioms must be asserted about probabilities in order that the calculus of probability may be applicable to them; but the difficulty is to relate these axioms to facts. How wide a difference of opinion still exists in this matter may be seen from any discussion between adherents and opponents of the so-called frequency theory of probability‡.

* The usual term "population" is undesirable here, for it has too definite connotations.

† I use *probability* to denote a property of propositions, and distinguish it from *chance*, which is a property of material systems. See N. R. Campbell, *Phil. Mag.* **44**, 67 (1922).

‡ See for instance the controversy between R. A. Fisher and H. Jeffreys, *Proc. roy. Soc.* A, **138**, 48 (1932); **139**, 343 (1933); **140**, 133 (1933); **146**, 1 (1934). The question which led to the controversy has been settled in the opinion of many qualified to judge; but its settlement leaves quite unreconciled the directly contradictory statements made concerning the basis of all statistical treatment.

However, there are collections of two sorts to which everyone admits statistics to be relevant; they are typified respectively by the urn containing black and white balls and the die or the tossed coin. The first sort of collection is finite and enumerable; it is possible for the proportion of members of the various kinds to be accurately known when the sample is drawn. The second sort is not finite or even definite; it is impossible that this proportion should ever be known accurately; but, on the other hand, very large samples reveal that there is one single proportion, or possibly a narrow range of proportions, definitely characteristic of the collection. Some would say that this proportion is the limit to which the proportion in the samples tends as their size is indefinitely decreased; others deny the validity of that conception and maintain that the proportion can be ascertained only with the help of *a priori* assumptions. But they all seem to agree that the large samples are necessary; by no other means can (e.g.) a true die be distinguished from a weighted die or from something that is not regulated by chance at all.

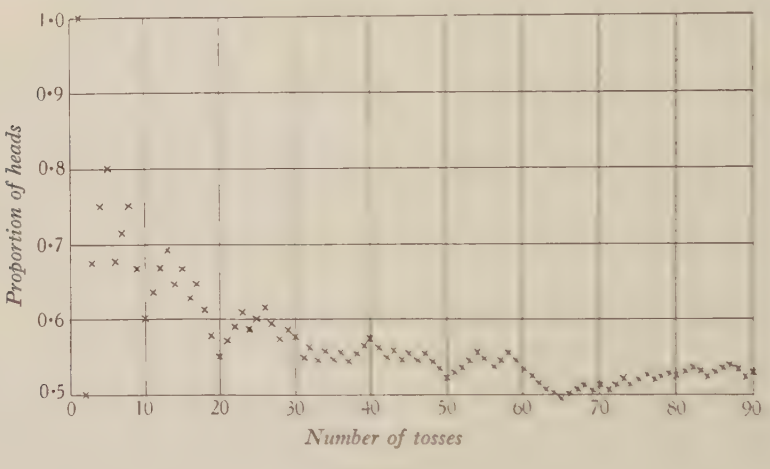


Figure 1.

The collection N_3 is clearly not of the first sort; it is not finite and enumerable. It may be of the second sort; but then we have no large sample. It must be insisted with the greatest possible emphasis that a single sample in which no regularity is evident does not afford sufficient evidence that the collection from which it is drawn is of the second sort and possesses the necessary characteristic proportion, or chance. On this point all serious statisticians seem agreed.

It is unusual to have as many as 20 measurements made in conditions so similar that, on any reasonable theory of error, they can be regarded as a sample drawn from a single collection. How inadequate such a sample is to prove the existence of a chance is illustrated by figures 1 and 2. These show the results of 500 tosses of a coin recorded for this occasion; the proportion of heads is plotted against the number of trials in the two diagrams on different scales. The first 20 tosses suggest no characteristic chance; the first 90 suggest one definitely greater than 0·5; the

first 500 leave the whole matter doubtful. Many more tosses would be required to convince anyone that the chance really exists.

The application of statistics to adjustment cannot therefore be justified by a similarity between the measurements to be adjusted and a collection of either of these two typical sorts. Many statisticians would doubtless maintain vehemently that their methods are applicable to other sorts of collections and samples; but the arguments advanced for such application are very various and often mutually inconsistent. In particular it seems to me that the arguments for applying the conception of probability to other sorts of collections are often destructive of statistics in the sense in which I am using the word, namely the rules advanced by Deming and

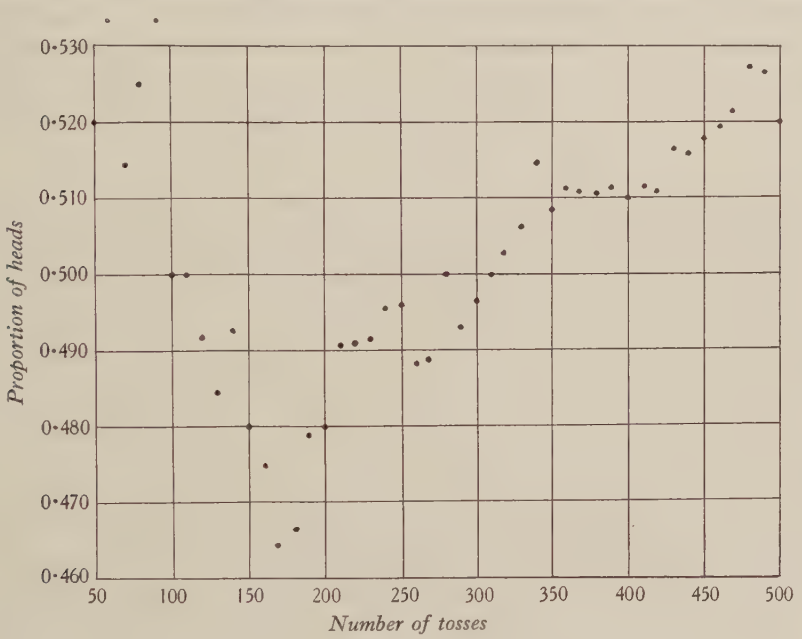


Figure 2.

Birge for reaching conclusions concerning a collection from a sample of it. If those rules are valid in respect of our problem, no generally accepted proof of their validity has yet been offered.*

Having thus found that (in terms of the analogy) it is by no means certain that the game that physicists are playing is statistics, let us inquire whether there is any other game, having definite rules, according to which their procedure can be justified. In the next section I shall try to state such rules.

* Those who think that it is easy to decide when and why statistical theory *is* applicable might study this question. Complete rainfall records are available for the years 1901–20; was the climate between 1901 and 1910 different from that between 1911 and 1920? If this problem is statistical, the records for 1901–10 and for 1911–20 must be samples of two possibly different collections. But what are the remainders of these collections? Not the records for other years; for, if the climate may be changing, other years are not comparable. But meteorological records must be records for some defined period. If the records for 1901–10 are a mere sample of the records for some longer period, and not the whole collection relevant to the problem, what is this longer period?

§ 3. THE ADJUSTMENT OF OBSERVATIONS

x *The simplest problem.* Suppose that I am measuring a well-defined length x on a millimetre scale. I put one end against the zero and denote the position of the other X_1 end by a numeral X_1. Subsequently I repeat the process and record another X_2 numeral X_2. If X_1 is the same as X_2, no question of adjustment can arise. If X_1 and X_2 are different, there are two alternative assumptions. Either the length has changed, or one (or both) of the observations are in error. If the first alternative is adopted there is no justification for any adjustment, apart from certain circumstances to be discussed on page 806. I have measured two different things, and there is an end of it. Only the second needs examination.

The reason why we are prepared to consider the second alternative is this. When I am assigning X_1 or X_2, I am never quite sure what numeral I ought to assign. All that I can really do, at any observation of any magnitude except number, is to α, β distinguish two ranges for X, namely a wide range α and a narrow range β within it. α is the range inside which I am sure that x must lie. If any subsequent observation lies outside α, then either the length has changed, or I have simply made a mistake, such as I might make even in counting. Mistakes can always be eliminated with sufficient care; we shall always assume here that they are eliminated; no observations in which they may occur ought to be discussed at all. β is the range within which I have no preference for one numeral rather than another; it is conveniently expressed in decimal notation by limiting the number of places. Thus if X_1 is 56·3, I am only asserting that the range α is (say) from 56·1 to 56·5, and the range β from (say) 56·25 to 56·35. In virtue of this uncertainty in assigning any X, two values X_1 and X_2 are not necessarily inconsistent. If their β ranges overlap, they are as consistent as I can possibly expect; if their α ranges overlap, these are possibly consistent; they are inconsistent only if their α ranges lie outside one another.

That is the only primary and essential source of error. Once we realize the nature of error, the problem of adjustment becomes simple and its solution obvious. If the β ranges of all our observations overlap, the assignment to x of any value lying in their common part will assert a scientific fact as definite as any fact can be. The selection of one of these values rather than another (e.g. the mean or median) is purely conventional and adds nothing to the fact. If the α ranges but not the β ranges overlap, then there is no definite scientific fact that can be asserted with full confidence; the possibility cannot be excluded that the magnitude is really different at different observations and therefore that adjustment is not permissible. If we decide to reject this possibility, there is some reason to accept the value that lies farthest from the boundary of any α range; for the subjective probability of a variation of the magnitude is greater, the closer is any value that we accept to the boundary of its α range. This preferred value usually approximates closely to the mean of the observations; there is then some reason for preferring that value, but even more for preferring the median.

These considerations determine also how many observations I should make. If I can get a few observations with overlapping β ranges, no increase of their number

can possibly increase the certainty of any value that may be selected. If all further observations still overlap in their β ranges the field of selection may be narrowed, but not the certainty of the scientific fact. But if any observation should occur which overlaps only in α ranges, doubt is immediately introduced; if eventually a condition were reached in which only a very narrow range common to all α ranges survived, the range of selection would be very narrow; but the evidence for any value in that range would be practically nil. Accordingly no more observations should be taken than are sufficient to establish that values with overlapping ranges, α or β, can be obtained.

Curve-fitting. But pure science is never concerned with the value of a single magnitude; it is interested only in relations between magnitudes. The problem that we have just discussed is therefore of little importance, but the discussion enables us to deal briefly with important problems. Of these one will be taken as typical; the application of our treatment to the remainder will be obvious*.

In this problem we desire to know whether a numerical law

$$f(x, y) = a \qquad \text{.....(1)}$$

is true. x, y are two different magnitudes, f is a known function defining the form of the law, a is a constant whose value is unknown. We measure n examples x_r ($r = 1$ to n) of the magnitude x, and the magnitude y_r corresponding to each of them. (The exact meaning of "corresponding" need not be discussed here.) For x_r we shall obtain a set of observed values X_{rs} ($s = 1$ to m_r) and for each y_r a set $Y_{rs'}$ ($s' = 1$ to m_r'). We try to find in the β ranges of each X_{rs} and each $Y_{rs'}$ values (X_{ro}, Y_{ro}), and also a value of a, such that

$$f(X_{ro}, Y_{ro}) = a \qquad \text{.....(2)}$$

for all values of r. If we succeed, then our previous discussion shows that we have all the evidence for the truth of equation (1) that we can have for any scientific fact. If we fail in this case but succeed when we use α ranges instead of β ranges, then the law is possible but not certain; it is the less probable the more we have to use values X_{ro}, Y_{ro} near the boundary of their α ranges. If we cannot find values in the α ranges, the law is certainly false.

The way in which this test is usually applied is slightly different in detail, but the same in principle. It is used because it is much more convenient, and indeed indispensable in all but the very simplest cases. We select a value of a by some rule or method, such as drawing a straight line, which requires only a knowledge of the X_{rs}, $Y_{rs'}$. We then find the residuals resulting from the substitution of the X_{rs}, $Y_{rs'}$ in equation (1), and compare these residuals with certain ranges calculated from the α and β ranges of the observations; it is unnecessary to set out here the well-known process of calculation. The three conclusions then follow according as the residuals all fall in the ranges calculated from the β ranges, in those calculated from the α ranges, or outside both.

The conclusions thus attained apply only to the law with the selected value of a. If the law does not turn out to be certainly true we ought, in all strictness, to repeat

* Thus a relation between universal constants can be considered as a numerical law of which there is only one example.

the test with other values of a. But there is one method of selection (or curve-fitting) of which it is commonly assumed that, if it does not give a value of a which makes the law true, no such value exists; this is the method of least squares.* I do not desire to throw any doubt on this assumption, which I believe to be true, but it must be insisted that the formal evidence for it is very slight; I have never seen a deliberate test made. The further assumption, often made, that the method of least squares is the only method of which this is true, and therefore that the value of a obtained by it is the only "right" value, rests on even less evidence. Indeed such evidence as exists is directly contrary; examples can easily be found in which the value obtained by other methods is so close to that obtained by the method of least squares that, if one is satisfactory, there can be no evidence against the other.

These alternative methods have been proposed for the smaller amount of computation they involve and the simplicity of that which remains. The simplest is that known as zero sum; some others require even less arithmetic, but are a little more difficult to grasp in the first instance.†

Some of them are professedly, others actually, valid only in a limited class of cases; some selection among them may be needed. The advantage of the method of least squares—probably the only advantage—is that it can be legitimately applied without examination to any problem in which curve-fitting is permissible at all, that is to say, in which the form of the law given by f is definitely known.

Corrections. The very simple conditions discussed on page 805 hardly ever arise in practice; usually corrections have to be applied to the directly observed magnitude. Thus in measuring length, a correction for varying temperature may have to be applied. Now it is definitely known that we are not, or may not be, measuring the same magnitude at each observation; but adjustment is still permissible for the following reason. In applying the correction we are in effect assuming that there is some law

$$\xi = F(x, u, v, \ldots, a, b, c, \ldots) \qquad \ldots\ldots(2),$$

x, u, v a, b, c ξ

where x is the magnitude actually observed, u, v, ... are other measurable magnitudes in virtue of which corrections must be applied, a, b, c, ... are known constants, and ξ is the magnitude that we really want to determine. (Thus x may be measured length, u temperature, a a coefficient of expansion, b a standard temperature and ξ a length at the standard temperature.) We measure x, u, v, ... and calculate ξ from equation (2). The measurements of x, u, v, ... will all have their α and β ranges, unless one of them is number; and corresponding to their α and β ranges (in a sense that again need not be explained) will be α and β ranges for ξ. If we use these calculated ranges for the corrected magnitude, instead of the original ranges for the uncorrected magnitude, all our discussion of tests proceeds as before. The tests again are, in principle but not exactly in detail, those actually applied by the best experimenters in their most careful work.

Sometimes, perhaps usually, in addition to these known corrections there are

* By the "method of least squares" I mean an empirical rule for selecting one out of an infinitude of approximate solutions of a set of equations. I imply nothing about the theoretical assumptions by which the use of this rule is sometimes justified.

† See, for example, T. Smith, *Proc. phys. Soc.* **46**, 560 (1934) and J. H. Awbery, *ibid.* p. 574.

unknown corrections. Thus we may know that the temperature of the rod is not uniform, but be unable to determine the distribution of temperature. If the error due to neglecting these corrections is not known to be limited, there is no justification for making any measurements at all. If it is known to be limited and the limit is known, the effect of neglecting the correction is to extend the α range. But there will not be in general anything corresponding to a β range. If the error is known to be systematic, i.e. always of one sign, there is certainly no β range; and even if there is no reason to suspect it to be of one sign rather than another, and if there is a range of complete and uniform uncertainty, it has not the same qualities as a true β range. For a true β range is inherent in the nature of measurement and of numerical laws; if there were not β ranges, no numerical law could be true.* An uncertainty due to an ignored correction is not of this character.

If then there are ignored corrections giving rise to an error comparable with the β range of the measurement, there can never be certainty either of an adopted value or of a numerical law. For only α ranges are present and any test based on them can never establish a certainty. In these circumstances it is best not to make any adjustment for the ignored correction. We should adjust as if it were absent, and then merely point out what degree of uncertainty attends the result.

§ 4. CONCLUSIONS

Such, I suggest, are the rules of adjusting observations as practised by the best experimenters.† There is necessarily some uncertainty attending the process. Accordingly those who believe that the calculus of probability is applicable to all propositions that are neither certainly true nor certainly false will hold that the calculus ought to be applied to adjustment. But here I am asking only whether the particular branch of the calculus concerned with collections and their samples is applicable; let us examine this question again, directing our attention now to condition (*b*) of page 801.

There are three kinds of collection which are, or may be, concerned in adjustment. First, if we are examining the truth of a law there is the collection of all examples of the law, which defines its form. Knowledge of this collection would be useful, but statistics does not profess to give it. The form of the law is always assumed to be known, and the problem is thus made substantially identical with that of determining a single value.

Secondly, there is the collection of all possible unknown corrections. But knowledge of this does not help us. We want to find out, not what corrections might be necessary (for we usually know that), but what corrections are necessary; in other words it is the sample, not the collection, that interests us. Further, it should be observed that, when (as is usual) the unknown corrections tend in one direction rather than the other, statisticians admit that they are powerless; for this is the case of systematic error.

* See, for example, my *Measurement and Calculation*, p. 153.

† As an example of this practice the paper by J. E. Sears, Jr., and H. Barrell, *Philos. Trans.* A, 233, 143, 216 (1934), may be quoted. But of course that is merely an example.

Thirdly, there is the collection of all values lying within the α (or β) range. Of this we do require information; for the collection defines the limits of the range. If our only source of information about the range were the recorded observations, then statistics would be useful—if it could determine the collection from the sample. But these are not the only source. The range arises from a certain indeterminacy in the judgment of the observer; we have, or should have, the testimony of the observer himself, which is the really decisive criterion. If the observer has been too careless or too incompetent to give us his opinion, his work is not worth considering at all. In any case it is surely the height of folly to deduce something about the judgment of an observer by means of an assumption that the observer himself would deny.

But such an assumption underlies all statistical treatment. Deming and Birge assume that "the effect of accidental errors can be reduced as far as desired by taking enough observations". If this means (as it must, if it is to be relevant) that the value of the measurements increases indefinitely with their number, the statement is one that every experimenter would deny. He limits the number of his observations, not because he is lazy, but because he knows that certainty is decreased, rather than increased, by increasing them beyond a quite small number. His reason for making more than a single observation is not to find out what is his α range; that he knows when he takes a single observation. It is to gain some assurance about unknown corrections, to ascertain that the actual spread of his observations is not so much greater than that arising from his α range that variable unknown corrections have to be postulated. He refuses to increase his observations beyond a certain point because he knows that thereby he increases enormously the probability of introducing new unknown corrections.

Herein lies the fundamental fallacy of the statistical theory of errors. It seeks to eliminate completely the judgment of the observer, which is an essential factor in determining the worth of his observations.* The plain fact is that if the observer is not capable of making the best of his measurements, nobody else can do it—except possibly by drawing his attention to mere mistakes that he will immediately recognize as mere mistakes.

To say this is not to deny great value to the work of those who collate the work of different experimenters. If they themselves are possessed, as Prof. Birge is, of

* The absurdity reaches a climax when the results of different observers working in entirely different conditions are treated as a sample of a single statistical collection. But this is not an essential error. Here I am avoiding criticism of the erroneous, but unessential, assumptions of statisticians; but it may be worth while to draw attention to what seems to be a new error.

Deming and Birge assert (*loc. cit.* p. 124) that "the step . . . of the instrument . . . has the effect of grouping the observations into class intervals". If we are observing a single magnitude there is no plurality of values which the observations may classify, as the observations of (say) human stature may classify the statures of the individuals composing the population. Such classification by the measuring instrument in virtue of its "step" can arise only if the magnitude under observation is varying during observation. It may be—I think it is—meaningless to assert that there is a "true" value, and therefore meaningless to assert that the true value is constant; but then it is equally meaningless to assert, as apparently Deming and Birge do, that the true value is varying and assuming a plurality of values which the instrument classifies. In other words, there are involved in a measurement at most two things, the object we are trying to measure and our measurements of it. Deming and Birge seem to assume that there is a third thing intermediate between them; I cannot see what this third thing can be.

experimental insight, they may well reach valid conclusions which each of these experimenters has missed. But the value of their work lies in their experimental insight and not in their statistical technique. If the elaborate methods of computation that Prof. Birge uses in his discussion of universal constants are replaced by the simple methods used by every physicist, while his judgments concerning the merits and interrelations of the various determinations are retained, everything of real value will survive. A neat series of best values, each with a tidy probable error, will not be obtained; but this series is positively misleading; it suggests something contrary to every principle of measurement, namely that there is a "true" value that can be found by experiments on a single magnitude. But his conclusions concerning the relations between the universal constants, such as that predicted by Eddington, will still be reached. Let anyone who doubts that carry through the work for himself, using nothing but arithmetic means and graphical curve-fitting.

REFERENCE

(1) Deming, W. E. and Birge, R. T. *Rev. Mod. Phys.* **6**, 119 (1934).

621.385.2

THE VARIATION OF VOLTAGE-DISTRIBUTION AND OF ELECTRON TRANSIT-TIME WITH CURRENT IN THE PLANAR DIODE

By R. COCKBURN, B.Sc.

Communicated by Dr D. Owen, February 28, 1935. *Read in title June* 7, 1935

ABSTRACT. The general solution for the voltage-distribution in a diode with plane electrodes is derived from the differential equation of motion of the electrons, the effect of initial velocities being treated approximately. The solution is applicable for currents from zero up to that value which reduces the cathode field to zero. The general solution for electron transit-time is then derived. It is found that increase of space charge increases the transit-time, most of this increase occurring just before the cathode field is reduced to zero. The effect of initial velocities of emission is to reduce the transit-time, the amount of the reduction increasing as space charge increases, so that the sharp increase of transit-time due to the accumulation of space charge becomes less apparent for large initial velocities. The application of the theory to electron oscillations is discussed.

§ 1. INTRODUCTORY

In the production of electron oscillations by means of the thermionic triode valve, it is generally accepted that the periodic time of these oscillations is dependent on the transit-time of the electrons between the valve electrodes. A method of calculating the variation of transit-time with current in the grid-anode space, after saturation of the space has occurred, has been given by Gill[1], but the variation with current of the transit-time in the cathode-grid space has not previously been treated. In the present paper the voltage-distribution and the electron transit-time between plane electrodes have been worked out for any value of current between the electrodes, excluding the case when a potential-minimum is formed in front of the cathode. The effect of initial velocities of emission, which cannot be neglected, has been taken into account.

The problem of the transmission of electrons between plane electrodes was first examined by Langmuir[2], who obtained the differential equation of motion of the electrons. Child[3] had previously obtained the same equation and applied it to the motion of positive particles. Langmuir neglected initial velocities and gave a particular solution for the voltage-distribution applicable only to the one value of current which reduced the field at the cathode to zero. In a later paper Langmuir[4] obtained the voltage-distribution taking into account the initial velocity-distribution of the electrons. The rigorous solution which he gave is only possible when the current between the electrodes is sufficiently great to produce a potential-minimum in front of the cathode.

§ 2. INITIAL VELOCITIES OF EMISSION

To obtain a solution applicable for all values of current from zero up to that value which reduces the field at the cathode to zero, it has been necessary to make an assumption with regard to the initial velocities of the electrons emitted by the cathode. Whereas according to Maxwell's law they are emitted with all velocities from zero to infinity, it is here assumed that they are all emitted with an average velocity u_0 such that u_0

$$u_0 = \left(\frac{\pi kT}{2m}\right)^{\frac{1}{2}},$$

where T is the absolute temperature, k is Boltzmann's constant, and m is the mass of the electron. $T, k,$

The error thus introduced in calculating the voltage-distribution is zero when the current flow is small, since in this case the voltage-distribution is independent of initial velocities; but as the current increases the error will become sensible. In the limiting case in which the current is sufficient to reduce the cathode field to zero, the approximate solution may be compared with Langmuir's rigorous solution[4]. The error is found to be normally of the order of only 1 per cent; see equation (14*a*) *et seq.* For greater currents, for which the cathode field is negative and a potential-minimum is formed in front of the cathode, this assumption with regard to initial velocities is not valid, since an initial velocity-distribution is implicit in such a condition.

Likewise in the calculation of the transit-time, the equation obtained gives in effect the transit-time of an average electron. This differs, however, from the average transit-time by an amount which is negligible in practice.

§ 3. THE VOLTAGE-DISTRIBUTION

Let anode and cathode be assumed to be infinite parallel planes separated a distance x_a, and let the cathode be capable of emitting any given number of electrons at an average velocity u_0. Let the anode be maintained at a constant positive potential V_a, the cathode being at zero potential. Let the current-density be uniform and equal to i. x_a V_a i

At a distance x from the cathode, let V be the potential, u the velocity of the electrons, n the number of electrons per cm^3, and κ the dielectric constant of the medium and let m and e have their usual meanings. V, u n, κ m, e

Then, following Langmuir, we may write

$$i = neu \qquad \ldots\ldots(1),$$

$$\tfrac{1}{2}mu_0^2 + Ve = \tfrac{1}{2}mu^2 \qquad \ldots\ldots(2),$$

$$\frac{d^2V}{dx^2} = \frac{4\pi}{\kappa} ne \qquad \ldots\ldots(3).$$

From equation (2)

$$u = \sqrt{\frac{2e}{m}} \cdot (V + V_0)^{\frac{1}{2}},$$

V_0 where
$$V_0 e \equiv \tfrac{1}{2} m u_0^2 \qquad \text{.....(4).}$$

On substitution in equation (1)

$$ne = \sqrt{\frac{m}{2e}} \cdot i\,(V + V_0)^{-\frac{1}{2}};$$

then from equation (3)

$$\frac{d^2V}{dx^2} = \frac{4\pi}{\kappa} \sqrt{\frac{m}{2e}} \cdot i\,(V + V_0)^{-\frac{1}{2}} \qquad \text{.....(5).}$$

The assumptions $V_0 = 0$ and $dV/dx_{x=0} = 0$ give Langmuir's[2] particular solution of the equation, which may be put in the form

i_s
$$i = i_s = \frac{\kappa}{9\pi} \sqrt{\frac{2e}{m}} \cdot \frac{V_a^{\frac{3}{2}}}{x_a^2} \qquad \text{.....(6).}$$

To obtain the general solution we multiply both sides of equation (5) by dV/dx and integrate. Thus

k
$$\frac{dV}{dx} = \{4\alpha\,(V + V_0)^{\frac{1}{2}} + k\}^{\frac{1}{2}} \qquad \text{.....(7),}$$

α where
$$\alpha \equiv \left(\frac{4\pi}{\kappa} \sqrt{\frac{m}{2e}} \cdot i\right) \qquad \text{.....(8).}$$

To perform the second integration, put

y
$$y^2 = \{4\alpha\,(V + V_0)^{\frac{1}{2}} + k\};$$

A then
$$x = \frac{1}{4\alpha^2} \left\{ y \left(\frac{y^2}{3} - k \right) + A \right\}$$

$$= \frac{1}{4\alpha^2} \left[\{4\alpha\,(V + V_0)^{\frac{1}{2}} + k\}^{\frac{1}{2}} \left\{ \tfrac{4}{3}\alpha\,(V + V_0)^{\frac{1}{2}} - \tfrac{2}{3}k \right\} + A \right] \qquad \text{.....(9).}$$

Now $x = 0$ when $V = 0$,

$$\therefore A = -\{4\alpha V_0^{\frac{1}{2}} + k\}^{\frac{1}{2}} \left\{ \tfrac{4}{3}\alpha V_0^{\frac{1}{2}} - \tfrac{2}{3}k \right\}.$$

By means of equations (6) and (8) express α in terms of i/i_s, thus

$$\alpha = \frac{4}{9} \frac{i}{i_s} \frac{V_a^{\frac{3}{2}}}{x_a^2} \qquad \text{.....(10).}$$

k_1 The subsequent working is simplified by introducing the new constant k_1 defined by the relation

$$k = \left(\frac{4}{3} \frac{V_a}{x_a}\right)^2 \left\{ k_1^2 - \frac{i}{i_s} \left(\frac{V_0}{V_a} \right)^{\frac{1}{2}} \right\} \qquad \text{.....(11).}$$

Substituting these values in equation (9), we have

$$\frac{x}{x_a} \cdot \left(\frac{i}{i_s}\right)^2 = \left[\left(\frac{i}{i_s}\right) \left\{ \left(\frac{V + V_0}{V_a} \right)^{\frac{1}{2}} - \left(\frac{V_0}{V_a} \right)^{\frac{1}{2}} \right\} + k_1^2 \right]^{\frac{1}{2}} \left[\left(\frac{i}{i_s}\right) \left\{ \left(\frac{V + V_0}{V_a} \right)^{\frac{1}{2}} + 2 \left(\frac{V_0}{V_a} \right)^{\frac{1}{2}} \right\} - 2k_1^2 \right]$$

$$- k_1 \left\{ 3 \left(\frac{i}{i_s}\right) \left(\frac{V_0}{V_a} \right)^{\frac{1}{2}} - 2k_1^2 \right\} \qquad \text{.....(12).}$$

Equation (7) now becomes

$$\frac{dV}{dx}=\frac{4}{3}\frac{V_a}{x_a}\left[\left(\frac{i}{i_s}\right)\left\{\left(\frac{V+V_0}{V_a}\right)^{\frac{1}{2}}-\left(\frac{V_0}{V_a}\right)^{\frac{1}{2}}\right\}+k_1^2\right]^{\frac{1}{2}} \qquad \text{.....}(7a),$$

and when $V=0$ and $x=0$

$$\left(\frac{dV}{dx}\right)_{x=0}=\left(\frac{4}{3}\frac{V_a}{x_a}\right)k_1 \qquad \text{.....}(7b).$$

Thus k_1 is directly proportional to the field at the cathode.

Now since $V=V_a$ when $x=x_a$, k_1 must satisfy the following equation, obtained from equation (12)

$$\left(\frac{i}{i_s}\right)^2=\left[\left(\frac{i}{i_s}\right)\left\{\left(1+\frac{V_0}{V_a}\right)^{\frac{1}{2}}-\left(\frac{V_0}{V_a}\right)^{\frac{1}{2}}\right\}+k_1^2\right]^{\frac{1}{2}}\left[\left(\frac{i}{i_s}\right)\left\{\left(1+\frac{V_0}{V_a}\right)^{\frac{1}{2}}+2\left(\frac{V_0}{V_a}\right)^{\frac{1}{2}}\right\}-2k_1^2\right]$$
$$-k_1\left\{3\left(\frac{i}{i_s}\right)\left(\frac{V_0}{V_a}\right)^{\frac{1}{2}}-2k_1^2\right\} \qquad \text{.....}(13).$$

By rearranging and squaring we reduce this to

$$6\frac{i}{i_s}\left(\frac{V_0}{V_a}\right)^{\frac{1}{2}}k_1+3k_1^2-4k_1^3=\frac{i}{i_s}\left\{\left(1+\frac{V_0}{V_a}\right)^{\frac{3}{2}}+\left(\frac{V_0}{V_a}\right)^{\frac{1}{2}}\left(3-\frac{V_0}{V_a}\right)\right\}-\left(\frac{i}{i_s}\right)^2 \qquad \text{...}(13a).$$

The parameter equation (13*a*) gives three values of k_1 for each value of i/i_s. One positive root is introduced during the squaring operation and does not satisfy equation (13), and one of the other two roots is negative for all values of i/i_s. This is physically untenable since $dV/dx_{x=0}$ cannot be negative within the limits of the problem. There is thus a unique value of k_1 for each value of i/i_s and V_0/V_a.

Tables 1, 2 and 3 show the variation of k_1 with i/i_s for three different values of V_0/V_a. The given value of V_0/V_a is substituted in equation (13*a*) and the cubic in k_1 is then solved for the various values of i/i_s by successive approximations.

Table 1. Variation of k_1 with i/i_s; $V_0/V_a=0$.

i/i_s	k_1
0·0	0·7500
0·1	0·7047
0·2	0·6575
0·3	0·6080
0·4	0·5557
0·5	0·5000
0·6	0·4398
0·7	0·3733
0·8	0·2972
0·9	0·2028
1·0	0·0

Table 2. Variation of k_1 with i/i_s; $V_0/V_a=0{\cdot}001$.

i/i_s	k_1
0·0	0·7500
0·1	0·7071
0·2	0·6621
0·3	0·6154
0·4	0·5665
0·5	0·5155
0·6	0·4600
0·7	0·4005
0·8	0·3349
0·9	0·2598
1·0	0·1665
1·0963	0·0

Table 3. Variation of k_1 with i/i_s; $V_0/V_a=0{\cdot}0049$.

i/i_s	k_1
0·0	0·7500
0·1	0·7093
0·2	0·6673
0·3	0·6232
0·4	0·5784
0·5	0·5357
0·6	0·4812
0·7	0·4282
0·8	0·3713
0·9	0·3091
1·0	0·2389
1·1	0·1549
1·217	0·0

It will be seen that when $i/i_s \to 0$, $k_1 \to \frac{3}{4}$ whatever the value of V_0/V_a. By substituting these values in equation (12) and expanding the contents of the first

square bracket we obtain the relation between the potential V and the distance x from the cathode, in the case of a negligible current, as

$$V=\frac{V_a}{x_a}.x.$$

Again when $dV/dx_{x=0}=0$, $k_1=0$ from equation (7b), and thus from equation (13a)

$$\frac{i}{i_s}=\left(1+\frac{V_0}{V_a}\right)^{\frac{3}{2}}+\left(\frac{V_0}{V_a}\right)^{\frac{1}{2}}\left(3-\frac{V_0}{V_a}\right),$$

whence, on substitution for i_s, equation (6)

$$i=\frac{\kappa}{9\pi}\sqrt{\frac{2e}{m}}.\frac{V_a^{\frac{3}{2}}}{x_a^2}\left\{\left(1+\frac{V_0}{V_a}\right)^{\frac{3}{2}}+\left(\frac{V_0}{V_a}\right)^{\frac{1}{2}}\left(3-\frac{V_0}{V_a}\right)\right\} \qquad \text{.....(14)}$$

$$=\frac{\kappa}{9\pi}\sqrt{\frac{2e}{m}}.\frac{V_a^{\frac{3}{2}}}{x_a^2}\left\{1+3\left(\frac{V_0}{V_a}\right)^{\frac{1}{2}}\right\} \text{ approximately} \qquad \text{.....(14}a\text{).}$$

η Now $V_0=\frac{m}{2e}u_0^2=\frac{\pi}{4}.\frac{k}{e}.T$. If we put η equal to eV_a/kT, equation (14a) becomes

$$i=\frac{\kappa}{9\pi}\sqrt{\frac{2e}{m}}.\frac{V_a^{\frac{3}{2}}}{x_a^2}\left\{1+\frac{2\cdot66}{\sqrt{\eta}}\right\} \qquad \text{.....(14}b\text{).}$$

This coincides with the approximate solution given by Langmuir[4], who showed that the current given by this expression differs by less than 0·3 per cent from the exact value when $dV/dx_{x=0}=0$; for the small values of V_0/V_a normally obtained, the current given by equation (14) will accordingly differ by about 1 per cent from the exact value. Thus the assumption that all the electrons are emitted with the same velocity instead of with distributed velocities causes an error of the order 1 per cent in the limiting case when $dV/dx_{x=0}=0$.

§ 4. THE ELECTRON TRANSIT-TIME

From equation (2) we have

$$u=\sqrt{\frac{2e}{m}}.(V+V_0)^{\frac{1}{2}},$$

if we write $V_0e=\frac{1}{2}mu_0^2$ as before.

T Now the transit-time $T=\int_0^{x_a}\frac{dx}{u}$

$$=\int_0^{x_a,V_a}\sqrt{\frac{m}{2e}}.(V+V_0)^{-\frac{1}{2}}\,dx$$

and from equation (7) $$dx=\frac{dV}{\{4\alpha(V+V_0)^{\frac{1}{2}}+k\}^{\frac{1}{2}}};$$

whence integrating and writing as before

$$\alpha=\frac{4}{9}\frac{i}{i_s}\frac{V_a^{\frac{3}{2}}}{x_a^2}, \quad \text{and} \quad k=\left(\frac{4}{3}\frac{V_a}{x_a}\right)^2\left\{k_1^2-\frac{i}{i_s}\left(\frac{V_0}{V_a}\right)^{\frac{1}{2}}\right\},$$

we have

$$T=3\sqrt{\frac{m}{2e}}.\frac{x_a}{V_a^{\frac{1}{2}}}\frac{i_s}{i}\left[\left(\frac{i}{i_s}\right)\left\{\left(1+\frac{V_0}{V_a}\right)^{\frac{1}{2}}-\left(\frac{V_0}{V_a}\right)^{\frac{1}{2}}\right\}+k_1^2\right]^{\frac{1}{2}}-k_1\Bigg] \qquad \text{.....(15).}$$

It will be seen that when $i/i_s \to 0$ and $k_1 \to \frac{3}{4}$

$$T = T_0 = 2\sqrt{\frac{m}{2e}} \cdot \frac{x_a}{V_a^{\frac{1}{2}}} \left\{ \left(1 + \frac{V_0}{V_a}\right)^{\frac{1}{2}} - \left(\frac{V_0}{V_a}\right)^{\frac{1}{2}} \right\} \qquad \text{.....(16),} \qquad T_0$$

and when $k_1 = 0$ and $\dfrac{i}{i_s} = \left\{ \left(1 + \dfrac{V_0}{V_a}\right)^{\frac{3}{2}} + \left(\dfrac{V_0}{V_a}\right)^{\frac{1}{2}} \left(3 - \dfrac{V_0}{V_a}\right) \right\}$

$$T = T_s = 3\sqrt{\frac{m}{2e}} \cdot \frac{x_a}{V_a^{\frac{1}{2}}} \frac{1}{(1 + V_0/V_a)^{\frac{1}{2}} + 2\,(V_0/V_a)^{\frac{1}{2}}} \qquad \text{.....(17).} \qquad T_s$$

Equation (16) shows the effect of initial velocities on the transit-time when the current is negligibly small, and equation (17) shows the effect of initial velocities

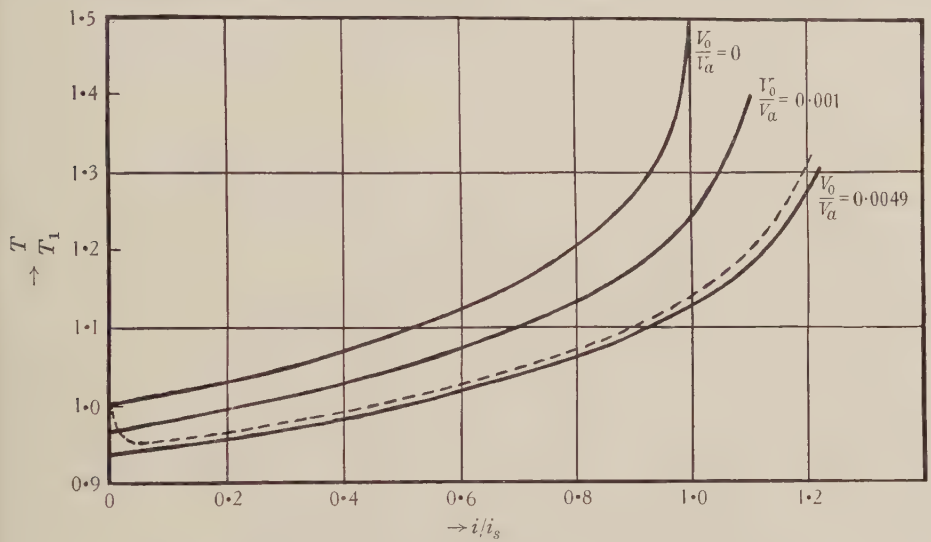


Figure 1. Variation of transit-time with current.

when space charge has reduced the field at the cathode to zero. Equation (15) with the parameter equation (13*a*) gives the value of the transit-time for any value of current, from zero up to that value which reduces the cathode field to zero, and for any value of emission velocity. It includes equations (16) and (17), of course, as special cases.

Figure 1 shows the calculated variation of transit-time T with degree of saturation i/i_s of the diode for different values of V_0/V_a. The ordinates are values of T/T_1 where $T_1 = 2\sqrt{(m/2e)}.x_a/V_a^{\frac{1}{2}}$. T_1

§ 5. APPLICATION TO EXPERIMENTAL CONDITIONS

In practice V_0/V_a varies with i/i_s, since both vary with the cathode-temperature, the emission velocity V_0 directly, and the current i nearly exponentially. Hence as i is increased from zero by raising the cathode-temperature, V_0 will increase rapidly at first and then quite slowly until $dV/dx_{x=0} = 0$. The current is then limited by space charge and a very large increase in temperature, and hence of V_0, produces only a small increase in i. The type of variation of V_0 with i to be expected is illustrated in figure 2.

These results were obtained from measurements on a valve of A.T. 40 type with a tungsten filament. The ratio of amperes of emission current to watts dissipated in such a filament varies only with its temperature, and the variation in the case of tungsten has been given by Stead [5]. From his results the approximate temperature of the valve filament for any voltage applied across it was obtained and the emission velocity was calculated.*

In figure 2 the abrupt change in slope to be expected at $dV/dx_{x=0}=0$ is rounded off owing to the velocity-distribution of the electrons, but it may be assumed that $dV/dx_{x=0}=0$ at point E, obtained by producing AB and DC to meet. Knowing now the value of V_0/V_a at E, we can calculate the value of i/i_s at this point from equation (14).

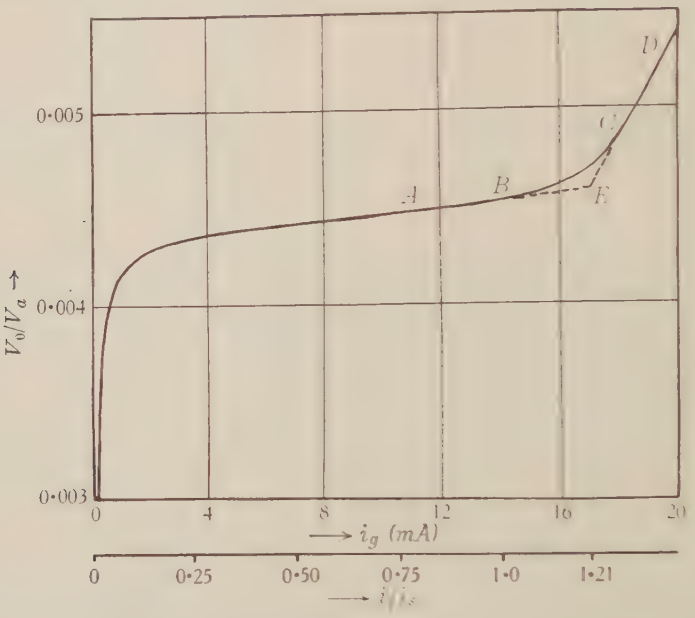


Figure 2. Variation of V_0/V_a with current for A.T. 40 valve at $V_a=50$ V.

The dotted curve in figure 1 was obtained by substituting in equation (15) for various values of i/i_s with the corresponding values of V_0/V_a from figure 2, and the corresponding values of k_1 from equation (13 *a*).

§ 6. CONCLUSIONS

As the current between the plane electrodes increases, the space charge being consequently increased, the transit-time of electrons between the electrodes also increases, and most of the increase occurs just before the cathode field is reduced to zero. The effect of initial velocities of emission is to reduce the transit-time, the amount of the reduction increasing as space charge increases, so that the sharp increase of transit-time due to the accumulation of space charge becomes less

* In the case of a cylindrical cathode $V_0=\frac{3}{2}\cdot\frac{\pi}{4}\cdot\frac{k}{e}\cdot T$.

apparent for large values of V_0/V_a. The effect of space charge on the electron transit-time should be most easily observed, therefore, with a low-temperature cathode and a large anode voltage.

The theory has been worked out for plane electrodes, whereas in practice valves with cylindrical electrodes are normally used. Neglecting the effect of initial velocities, Schiebe [6] has calculated the electron transit-time between cylindrical electrodes when space charge is negligible, and McPetrie [7] has dealt with the case when space charge has reduced the cathode field to zero. The increase of electron transit-time due to space charge is of the same order as for plane electrodes. In another paper McPetrie [8] has extended Schiebe's treatment to include the effect of initial velocities on the transit-time when space charge is negligible. For the usual internal cathode the effect is considerably less than for plane electrodes. A general solution for transit-time between cylindrical electrodes is probably impossible. It is reasonable to assume, however, that the results obtained for plane electrodes will apply qualitatively to cylindrical electrodes.

§ 7. APPLICATION TO THE THEORY OF ELECTRON OSCILLATIONS

In the course of an experimental investigation by the author into the electron oscillations produced by a cylindrical-electrode triode valve connected in a Barkhausen-Kurz [9] circuit, the variation of the periodic time of the oscillations with grid current is being examined. For valves with high-temperature tungsten cathodes the usual decrease of periodic time with increasing current is observed. This effect is usually explained as due to the variation of electron transit-time in the grid-anode space after it has become saturated with space charge. It may also be explained by assuming that the oscillations are those of electrons in statistical equilibrium between the electrodes, the decrease in periodic time being due to the increase of concentration of the electrons with current.

In the case of one valve, however, with a low-temperature oxide-coated cathode, the decrease in periodic time was followed by a small but quite definite increase as the cathode-grid space became saturated with space charge. This increase was most evident at high grid voltages. These preliminary observations indicate definitely that the periodic time of the electron oscillations is determined by electron transit-time and that the path of these electrons is partly in the cathode-grid space. The investigation is proceeding with a view to establishing the complete path of those electrons which determine the periodic time of the oscillations.

REFERENCES

(1) E. W. B. Gill. *Phil. Mag.* **12**, 843 (1931).
(2) I. Langmuir. *Phys. Rev.* **2**, 450 (1913).
(3) C. D. Child. *Phys. Rev.* **32**, 498 (1911).
(4) I. Langmuir. *Phys. Rev.* **21**, 419 (1923).
(5) G. O. Stead. *J. Instn elect. Engrs*, **58**, 107 (1920).
(6) A. Schiebe. *Ann. Phys.*, Lpz., **73**, 54 (1924).
(7) J. S. McPetrie. *Phil. Mag.* **16**, 284 (1933).
(8) J. S. McPetrie. *Phil. Mag.* **16**, 544 (1933).
(9) H. Barkhausen and K. Kurz. *Phys. Z.* **21**, 1 (1920).

541.133

THE ELECTRICAL CONDUCTIVITY OF SOME STRONG ELECTROLYTES IN DILUTE SOLUTION AND ITS VARIATION OVER THE TEMPERATURE RANGE 18° C. TO 85° C.

By C. J. B. CLEWS, B.Sc., A.Inst.P., Queen Mary College, University of London

Received February 28, 1935. *Read in title June* 7, 1935

ABSTRACT. The work described in a previous paper has been extended to the measurement of the electrical conductivity of 0·0005-N solutions of potassium nitrate, potassium sulphate and sodium sulphate. These data, in conjunction with those previously obtained, provide strong evidence for the validity of the theory of Debye, Hückel and Onsager for uni-univalent electrolytes in dilute solutions up to 85° C. Agreement is not as satisfactory for sodium sulphate, while an anomaly has been observed in the behaviour of potassium-sulphate solutions.

§ 1. INTRODUCTION

In a previous communication to the Society[1] some measurements of the electrical conductivity of aqueous solutions of potassium nitrate, potassium sulphate and sodium sulphate were reported and data were given for solutions having concentrations down to N 1000. Owing to the Debye-Hückel-Onsager theory being applicable only in very dilute solutions, no deductions could be made as to the agreement of experiment with this theory in the case of uni-bivalent electrolytes, and only tentative conclusions could be made for potassium-nitrate solutions. It was in view of this difficulty that the present measurements were made.

§ 2. EXPERIMENTAL

The conductance has been determined with the same apparatus as before, with the exception that a new variable audio-frequency oscillator was used.

In extending the range to the more dilute solutions many difficulties were encountered. Among these was contamination of the conductivity water due very largely to working in a dirt-laden city atmosphere. There was also some difficulty in making up such dilute solutions with the required accuracy.

Owing to these and other minor troubles, it was only possible to measure the conductance at one more concentration, namely, N 2000, for the three electrolytes.

§ 3. RESULTS

The further data are given in table 1. Comparison with the work of other experimenters cannot be extensive, but where it is possible, i.e. 18° and 25°, agreement is quite satisfactory.

Table 1. Equivalent conductance of $N/2000$ solutions

Temperature (° C.)	Potassium nitrate	Potassium sulphate	Sodium sulphate
18	124·4	128·5	107·5
25	142·5	150·2	125·0
30	156·8	167·8	139·6
35	171·2	187·5	153·2
40	186·6	204·8	169·3
45	202·3	224·1	185·5
50	218·0	242·3	203·2
55	235·5	260·8	217·8
60	251·9	278·5	234·3
65	268·0	297·7	253·1
70	284·8	318·1	268·8
75	302·0	337·4	287·6
80	318·8	356·3	307·3
85	335·5	375·6	323·7

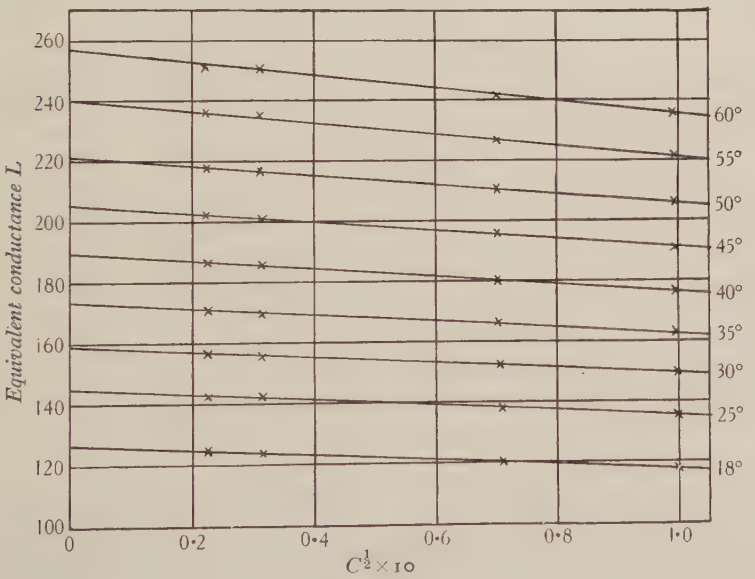


Figure 1. Potassium nitrate.

From the Debye-Hückel-Onsager theory we expect that a plot of the equivalent conductance against the square root of the concentration, at a given temperature, will yield a straight line; by this means we may readily test the theory. Figures 1

and 2 show how striking is the agreement for potassium-nitrate solutions, and furthermore, the values of the conductance at infinite dilution, L_0, given by the intercepts of these lines with the L axis are almost identical with those calculated from the theory, table 2*.

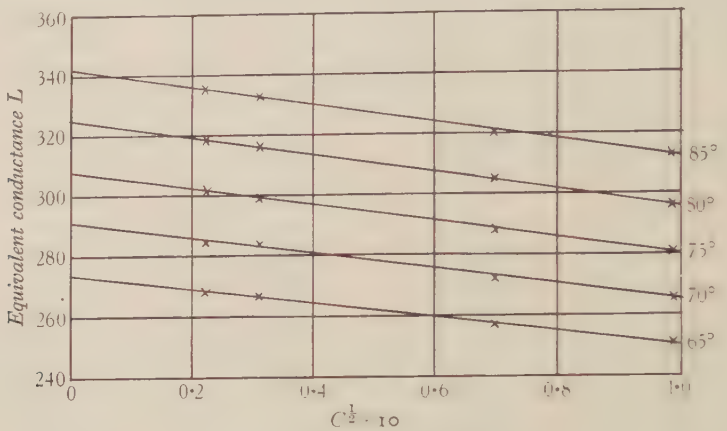


Figure 2. Potassium nitrate.

Table 2. Potassium nitrate. Equivalent conductance at infinite dilution

Temperature (° C.)	L_0 from graph	L_0 from Debye-Hückel-Onsager theory
18	126·3	—
25	144·6	144·54
30	159·1	159·04
35	173·8	173·78
40	189·7	189·68
45	205·6	205·39
50	221·5	221·30
55	239·5	—
60	257·1	—
65	274·2	—
70	291·3	—
75	308·1	—
80	325·7	—
85	342·4	—

It should be noted, however, that the experimental slope and the calculated slope $(\alpha L_0 + \beta)$† are not in absolute agreement; but the deviation is not serious, as may be seen from the observed and computed values of the conductance of an $N/100$ solution, table 3.

The results for potassium-sulphate solutions do not give us a linear graph, but give a curve which approximates to a parabola, except at the lower temperatures,

* Cf. reference (1), p. 770.

† α and β are factors involved in the Debye-Hückel-Onsager theory (*loc. cit.* p. 768).

figures 3 and 4. The curvature increases with the temperature. We are therefore led to the conclusion that even in fairly dilute solutions potassium sulphate does not behave as an ideal Debye electrolyte.

Table 3. Potassium nitrate

Temperature (° C.)	Debye slope	Experimental slope	L, calculated $N/100$	L, experimental $N/100$
25	92·9	89	135·3	135·7
30	103·8	97	148·7	149·4
35	115·2	112	162·3	162·6
40	127·3	131	177·0	176·6
45	139·6	145	191·6	191·1
50	152·0	154	206·3	206·1

Some explanation may be offered by the Bjerrum theory [2] of ionic association. For uni-bivalent electrolytes the critical distance of approach is probably of the order 7 A., whereas the sum of the ionic radii of potassium sulphate is 3·7 A., so that we might expect strong association. This is, however, nothing more than a tentative and qualitative explanation.

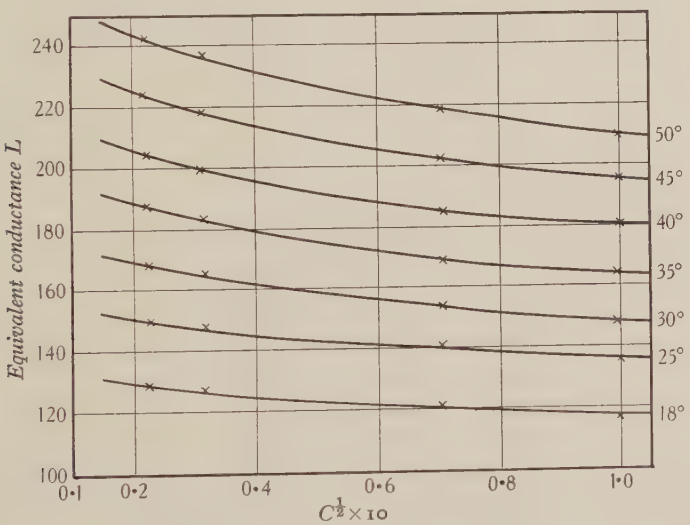


Figure 3. Potassium sulphate.

Contrary to expectations, sodium sulphate does not give parabolic curves; nevertheless, the linear plot is not as good as for the uni-univalent electrolyte.

The difference in the behaviour of these two sulphates provides a problem which may be rather difficult of solution, as it is usually found that similar groups of salts have very similar properties.

§ 4. CONCLUSION

The observations on potassium-nitrate solutions add definite support to the Debye-Hückel-Onsager theory, and particularly to its validity at higher temperatures. This is an extremely stringent test of the theory in that over the temperature-range 18° to 85° there is about a threefold increase in the conductance. We have already had ample proof, by Hartley [3] and his co-workers at Oxford, of the accordance of theory with experiment for dilute uni-univalent electrolytes at 25° in both aqueous and non-aqueous solvents. There can now be little doubt as to the fundamental truth of the Debye theory, particularly for uni-univalent electrolytes.

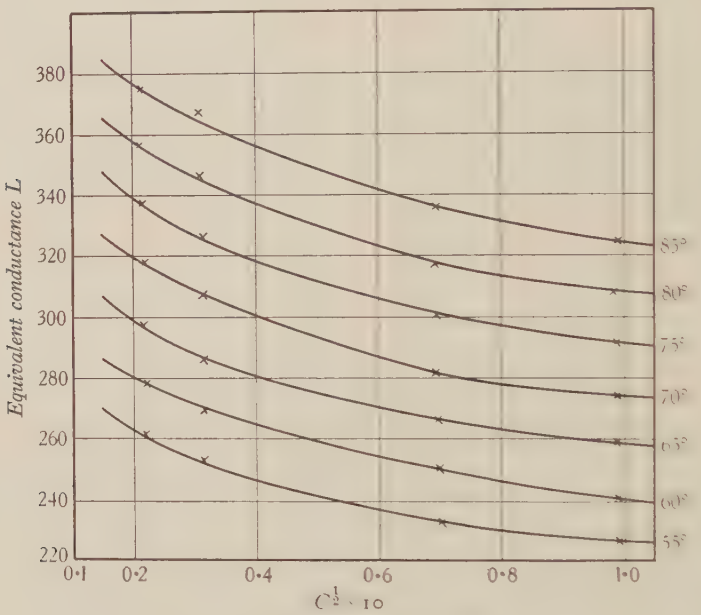


Figure 4. Potassium sulphate.

Rather less conclusive is the evidence offered by sodium-sulphate solutions, while solutions of potassium sulphate deviate from the theory even in dilute solutions. A tentative explanation is given by the Bjerrum association theory, but this is nothing more than qualitative.

That further information is required with regard to the apparently anomalous behaviour of potassium sulphate is obvious. Owing to local conditions the author will be unable to continue the investigations, and we can only hope that the necessary data, both on the behaviour of potassium-sulphate solutions and on the variation of conductance with temperature, may be forthcoming in the near future.

§ 5. ACKNOWLEDGMENTS

The work has been carried out in the Physical Laboratories at Queen Mary College, and my best thanks are tendered to Prof. H. R. Robinson, F.R.S., for the facilities he has placed at my disposal and for his advice throughout the investigation.

I am indebted to the Department of Scientific and Industrial Research for a grant which enabled me to carry out this work.

REFERENCES

(1) Clews, C. J. B. *Proc. phys. Soc.* **46**, 764 (1934).
(2) Bjerrum, D. *Kg. Dansk. Vid. Selsk. Medd.* **7**, No. 9 (1926).
(3) Hartley, H. *Proc. roy. Soc.*, A **109**, 351 (1925); **127**, 288 (1930); **132**, 427 (1931) etc.

537.56.06

AN IMPROVED COUNTING CIRCUIT

By H. P. BARASCH, Cavendish Laboratory, Cambridge

Communicated by C. D. Ellis, May 28, 1935. *Read June* 7, 1935

ABSTRACT. This paper deals with a simple Geiger-Müller counter circuit, the essential feature of which is the application of a special characteristic curve of a pentode valve. This $\{i_a, e_g\}$ curve has a saturation portion in the negative grid region. The output pulses from the circuit are equalized, free from secondary disturbances, and of shortened duration. By means of an appropriate resistance-capacity coupling to the thyratrons a back pulse is generated, so that the final pulse-shape is more suitable for triggering the thyratrons. A description of the simple pentode circuit is given and also of a universal set, consisting of four valves, which is suitable for coincidence counting. The special characteristic curve when applied to the coincidence circuit leads to improved discrimination between total and partial coincidences.

§ 1. INTRODUCTION

When a Geiger-Müller counter coupled to an amplifying circuit is in use irregularities are sometimes observed whose cause is to be found not in the nature of the discharge itself, but rather in the circuits between the counter and the recorder. Closer consideration proves that even with correct discharges in the counter, double pulses which affect the recording circuit may be generated in the amplifier.

A good amplifier should have the following qualities: (i) sufficient amplification to actuate the recording device; (ii) good resolution—here a distinction is to be made between normal counting of a few hundred a minute and high-speed counting of a few thousand a minute; (iii) limitation, by which is meant the equalization of the pulse-amplitude; (iv) a clean, uniform output pulse-shape to prevent disturbances and false counts. For coincidence work there are two further points, which must be considered, viz. the provision of (v) high resolving-power for coincidences of the different counters, and (vi) good discrimination between coincidences of n counters and coincidences of $(n-1)$ counters.

In the last few years many amplifying-circuits for Geiger-Müller counters have been described. They usually involve several stages of amplification of the resistance-capacity-coupled type, and for coincidence counting still other stages are added. For example, in the circuit devised by Johnson and Street[1] three stages are necessary to fulfil conditions (i), (ii) and (iii). The first two stages amplify and equalize the pulses, the third reverses them and shortens their duration. In this way good resolution for coincidence work is obtained but further amplification is necessary because in the shortening of the duration of the pulses their size has been reduced. Parallel chains, according to the number of counters used, are necessary for coincidence counting.

An amplifier of the Johnson and Street type undoubtedly satisfies the conditions specified, but in simpler circuits condition (iv) is rarely realized. This demands the generation of a sharp uniform pulse-shape without wobbling, and it is not easy to

revent secondary potential disturbances in the coupling elements due to variation n size and form of the discharge in the counter itself. In particular the shortening tage, because of its low input time-constant, may generate a rather strong back-ulse, which is amplified by the valve and may give rise to a second forward pulse in ne succeeding stages.

Visual observations on a cathode-ray oscillograph screen show that the output ulses from certain multistage amplifiers of the ordinary type not only are distorted nd of long duration, but are followed by a series of undesirable potential-variations.

This paper will describe a method of using a single pentode, which in itself atisfies all the necessary conditions as a three-stage amplifier and also produces a hape of output pulse particularly favourable for thyratron recording.

§ 2. COUNTER DISCHARGE

A comprehensive treatment of the action of a Geiger-Müller counter has recently een given by Werner[2]. Certain points in his arguments are relevant to the present roblem, since they enable us to see in a general way the form of the pulse which is

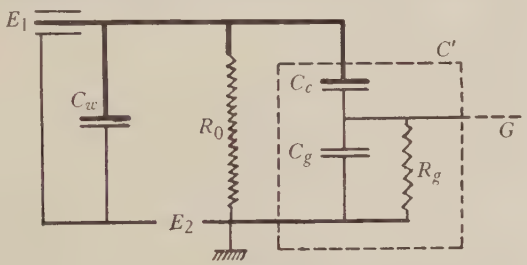


Figure 1.

ctually fed into the amplifying system, and how its shape and duration depend on he details of the coupling.

When an ionizing particle or quantum ionizes the gas in the counter, the anode ire acquires in a very short time some negative potential, depending on the over-oltage applied to the counter. There it remains in a region of unstable corona ischarge for some time t_m until the discharge stops abruptly. After this extinction, he wire returns to zero potential in a time determined by the net capacity and esistances associated with the counter.

The two factors, the time taken for the potential to rise and the time during hich the discharge burns in the counter, have to be considered in connection with he amplifying-circuit. One would expect that the former will determine the steep-ess of the pulse which arrives on the grid of the first valve. Werner has calculated he time t taken to charge the wire negatively to a small over-voltage and has ob-ained a value of the order of 2.10^{-7} seconds[2].

An equivalent scheme of the counter circuit is shown in figure 1. The counter is epresented by an e.m.f. E_1 with internal resistance K. C_w represents the wire-arth capacity including all direct connections. C_c is the coupling-condenser and C_g the effective grid capacity of the first valve with grid-leak resistance R_g. E_2

denotes the high-potential supply, assumed to have negligible internal resistanc and R_0 is the high resistance which serves to discharge the wire after extinctio K, the internal resistance of the counter, is a function of gas pressure, of the kind gas used, and of the geometrical counter constants. At small values of the ove voltage it may be obtained from formulae or from tables given in Werner's pape For a counter filled with air at a pressure of 8 cm. of mercury, K is of the order 1 megohm.

i, V' During the charging process, when a particle enters there is a current flo i equal to $-C_0 dV'/dt$, where V' is the over-voltage applied to the counter. If R_0 C_0 assumed to be so large that it does not affect the charging process, then C_0, the tot C' earth capacity of the wire, is given by $C_0 = C_w + C'$, where C' represents the equivale capacity of the impedance within the dotted rectangle. If we consider the pulse to b defined by the fundamental frequency ω we find that, approximately,

$$C' = \frac{C_c C_g (1 + 1/\omega^2 C_g^2 R_g^2)}{C_c + C_g (1 + 1/\omega^2 C_g^2 R_g^2)}.$$

Evaluating this formula with actual values of constants used and the fundament frequency as given by the slope of the primary charging current to the wire, we fin that $C_0 \cong 10$ to $15\,\mu\mu$F.

The time KC_0 of charging is correspondingly about 1.10^{-5} sec., a time muc longer than that obtained if no external load is connected to the counter, for whic we have already quoted Werner's estimate of 2.10^{-7} sec.

It is to be observed that in order to make the charging-time short it is essential the direct earth capacity C_w of the wire, shown by the dark lines in figure 1, th should be made as small as possible, whereas the coupling-capacity C_c has a rath indirect effect, because of its position in series with C_g and R_g. The very sm capacities involved render it necessary to avoid all superfluous metal parts connect to the counter wire which could increase the direct earth capacity. For this reason was found convenient to make the coupling-condenser C_c by enclosing an extensi of the counter wire in glass tubing of small diameter and wrapping metal foil arou the outside of the glass. The capacity between the wire and the foil can be ma about 20 to $40\,\mu\mu$F. without directly increasing the earth capacity.

In order that the wire potential after extinction shall come back to zero in a sho time, so that the next pulse falling within the recovery time $R_0 C_0$ is large enough f amplification, it is necessary to make R_0 small. But a small R_0 means that t negative rise of wire-potential, governed by K, will be partially offset by charge fro E_2 through R_0. This delays the rise of the wire-potential to the starting potenti value and also forces the discharge to proceed to burn in a region of instable coro discharge which is more remote from the actual starting potential. In this region is increased and the counter tends to continuous discharge. Too great a value of ov voltage has a similar effect. It is necessary therefore to have R_0 at least 100 times great as K. A compromise must be made between a correct working of the count and a reasonable resolving-power of the amplifier.

§ 3. PENTODE CHARACTERISTIC

The special pentode circuit we are discussing is shown in figure 2. The characteristic feature is that in series with the anode is a high resistance of about 250,000 Ω., so that the anode potential is about 3 V. The screen grid is kept at from 30 to 40 V. Under these conditions the characteristic shown in figure 3 (*a*) is obtained. It resembles the usual saturation curve of a hot emitter, but is unique in that the grid-voltage change required to go from cut-off to saturation of the anode current can be of the order of one volt. This range of voltage will be referred to henceforth as the *range of limitation* and denoted by L, figure 3 (*b*). L

The curve follows at first the usual slope of a pentode characteristic, but at a certain point it bends over very sharply into a horizontal line which extends from about -1 V. to high positive values. It is outside the scope of this paper to give an

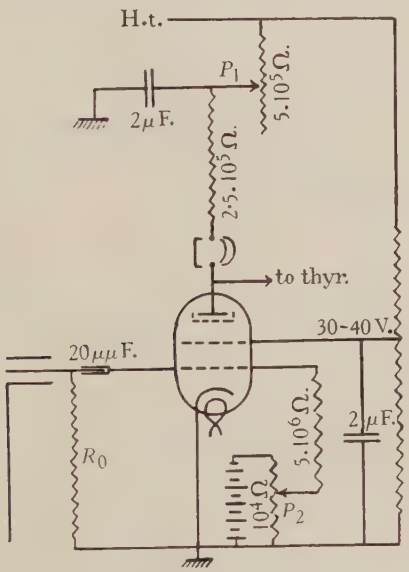


Figure 2.

exact physical explanation of this saturation portion, but it may be stated briefly, that it is due to a space charge being established near the plate when the plate potential is maintained somewhat below 5 V.[3].

In the present case it is the external resistance R_a which lowers the plate potential R_a
to about 3 V. A further consequence of importance is that the pentode valve in this region has an internal resistance of the very low order of about 5000 Ω. The behaviour in that state is analogous to that of a lighted thyratron, where the plate current is determined only by the anode resistance R_a and anode voltage V_a. The V_a
shape of the curve remains unchanged over a wide range of R_a and V_a. When R_a is decreased and V_a increased the current in the saturation part increases, but at the same time L increases. Satisfactory values found for a SP4 pentode valve are: h.t. voltage 150 V., $R_a = 250{,}000$ Ω. and screen voltage 30 to 40 V. The anode current i_a is then 0·63 mA. and $L = 1{\cdot}3$ V.

The working point A of the pentode is adjusted by means of a potentiometer to some point in the saturation region depending on the magnitude of the pulses on the grid. If this is sufficiently large, it is of advantage to shift A towards zero grid bias, because this gives an appreciable improvement in the shape of the pulse.

By the aid of this $\{i_a, e_g\}$ characteristic curve the conditions set out in the introduction are satisfactorily fulfilled. The adjustment in the saturation region corresponds to a bias which prevents not only the small backpulses but also all disturbances from being amplified; the latter may take the form of a.-c. and d.-c. hum, potential-variations due to thermal agitation in the high-input resistances, and general pick-

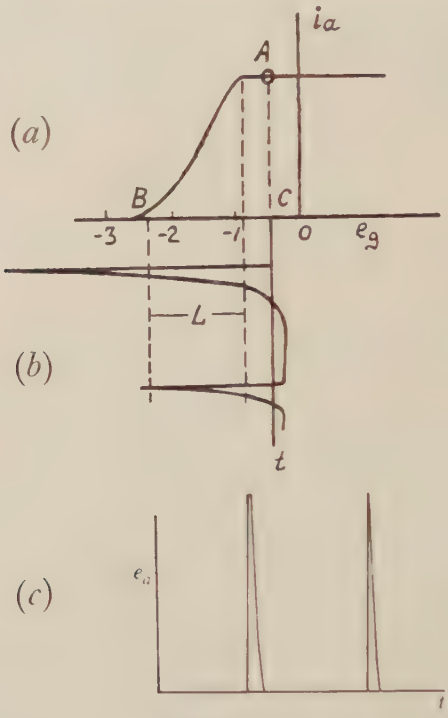


Figure 3.

up. Because of the small amplitude of these disturbances at the first grid, the bias (that is, the shift into the saturation region) need not be greater than a small fraction of a volt.

This pulse-cleaning effect of the biasing is also accompanied by a pulse-shortening action corresponding to the pulse-shortening stage of a multistage amplifier with low input time constants. This is due to the fact that only the steepest parts of a negative grid pulse are amplified, and the low-frequency discharge of the grid (of time constant $R_g C_g$) is cut off.

There is further a most important pulse-equalizing effect. To discuss these effects we will consider two pulses which, as shown in figure 3 (b), are separated by an interval of time less than the recovery time $R_0 C_0$ of the counter. Because of this the

second grid pulse is of smaller amplitude, but provided its peak extends beyond B in figure 3 (a) the output voltage as shown in figure 3 (c) will still be the same as for the first pulse. Further, the duration of the output pulses are much the same and become only slightly reduced when discharges fall close together.

The negative pulse impressed on the pentode grid consists of the negative rise (which, as we have seen, is of the order of 10^{-5} sec.) and of the return to zero potential, the time of which depends only on the grid time-constants. When $R_g = 5 \cdot 10^6\ \Omega$. and the grid capacity $C_g \cong 15\,\mu\mu$F., this return time may be taken as $7{\cdot}5 \cdot 10^{-5}$ sec., so that the duration of the whole pulse is a little less than 10^{-4} sec.

In order to examine the resolving-time for single counting with this amplifier we will refer to the diagram in figure 4. The curve R_0C_0 represents the recovery of the counter wire back to zero potential after a single count. As the time intervals between successive kicks decrease the input pulses decrease, and when they become less than BC, figure 3 (b), the amplitude of the output kicks becomes less than the normal equalized value. This is shown in figure 4, where the tips of the output pulses follow a curve such as f. A second pulse occurring after an interval less than t_l is not

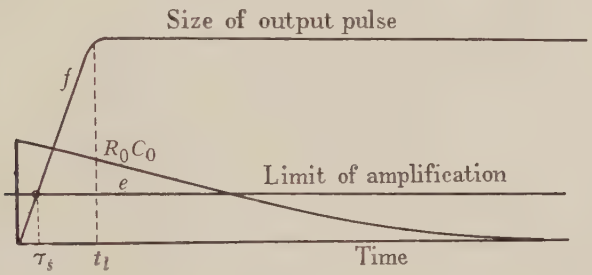


Figure 4.

equalized, because the corresponding grid pulse is smaller than BC, but it may still be recorded. If the $\{i_a, e_g\}$ characteristic curve were of a rectangular shape, the equalizing action could be made perfect. If the horizontal line e represents the minimum size of output pulses which will be recorded, the intersection of this line with the line f determines the resolving-time τ_s of the amplifier without thyratron and recording circuit. τ_s

It can therefore be seen that τ_s depends jointly on the value R_0C_0, on the pulse amplitude (which can be regulated by the adjustment of over-voltage) and on the amount of amplification. It may also be seen that as long as the pulses are single-peaked and of short duration their shape does not affect the resolving-power. The value of τ_s was measured by means of a pendulum breaking two contacts and found to be less than $\frac{1}{5000}$ sec. for the single-valve amplifier and about $\frac{1}{10000}$ sec. for the 3-stage amplifier mentioned later, even when the value of R_0C_0 was as high as $\frac{1}{10}$ sec.

In order to deal with fast counting and smaller pulses a universal amplifier with two fore stages was built, making use of this special pentode characteristic. A description of this amplifier is given in the next section.

§ 4. CIRCUIT

As has already been pointed out, in order to obtain this characteristic, only the values of R_a and the actual h.t. voltage V_a are of importance. The simple circuit has been shown in figure 2. V_a, and consequently the saturation current, can be varied by means of a 500,000 Ω. potentiometer P_1. The V_a point must be decoupled by means of a large condenser. On increase of the saturation current the horizontal part of the characteristic is simply raised, the slope of the inclined portion remaining the same, and the knee moving nearer to zero potential. There will be an optimum point of adjustment according to the size of the pulse from the counter. It appears experimentally that this adjustment is not critical, however, and in the simplest

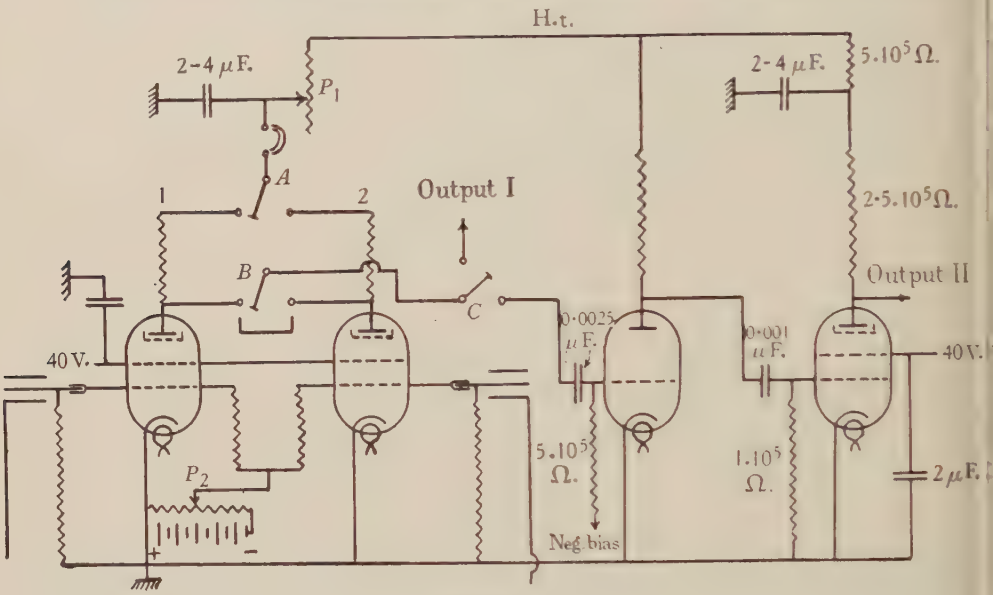


Figure 5.

case P_1 may be omitted and R_a may be connected directly to about 150 volts h.t. The potential of the screen grid is supplied from a voltage bridge across the source of h.t. and is kept steady by a 2 μF. decoupling condenser. The potentiometer P_2 is essential for the adjustment of the working-point. The most sensitive part of the characteristic is the upper knee of the curve, and when the starting potential of the counter is to be determined the working point is shifted to this region, which is easily detected by a.-c. hum in the telephones.

As has been outlined in § 1, care must be observed in mounting the components of the grid input circuit. To reduce all stray capacities the leads should be short and it is desirable where possible to have the counter located very close to the grid terminal. Since in this circuit the bias suppresses small disturbances, shielding of the input leads may be omitted, the input capacity to earth being thus reduced.

To deal with very small input pulses and to allow of coincidence counting, a universal amplifier, figure 5, which requires only 4 valves, was built. The first two

valves are SP4 pentodes arranged in parallel, each working exactly as in the simple circuit. The third valve serves purely as a voltage amplifier and may be any triode suitable for the purpose. An SPL with an amplification factor of 70 is satisfactory. The final valve is another SP4 pentode in a circuit differing from that for the input valves only by the omission of the two potentiometers P_1 and P_2. It is operated at zero grid bias. Switches A and B serve to adapt the amplifier for single counting by either of the input valves or for coincidence work. When the pulses are small switch C can be used to introduce the triode intermediate amplification and pentode final amplification.

§ 5. COINCIDENCE COUNTING

The adaptation of this circuit for coincidence incorporates the usual method developed by Rossi[4] and consists here of two equal SP4 valves. The working point A of each is slightly shifted into the saturation part of the characteristic.

As has been shown by Fussel and Johnson[5], pentodes are much superior to triodes for coincidence counting because the discrimination between partial and total coincidences is greatly improved. The form of the $\{i_a, e_g\}$ characteristic described in this paper seems to represent a further improvement. An amplifier using the normal pentode characteristic always shows small variations of output potential for partial coincidences. This effect is found to be reduced in the present amplifier for correct adjustment of the grid bias. The explanation is the fact that the pentode valve in the saturation region has such a low internal resistance that an anode pulse from the other valve is dropped to earth without affecting the output, while a coincidence input pulse of about 1 V. on the two grids causes a full anode swing.

To estimate the resolving-power of this amplifier for coincidences we note that pulses from the two counters will be separated if one pentode has returned to the saturation part, point A in figure 3, before the other starts to change. Now in § 3 we estimated that the duration of the whole change in the pentode was rather less than 10^{-4} sec., and we may take this figure as giving the order of magnitude of the resolving-time τ_c. A measurement of this quantity τ_c was carried out by use of the relation $\tau_c = D/2N_1N_2$, where D is the number of accidental coincidences between two counters when the number of discharges in each individually is N_1 and N_2. Values of τ_c from 5 to 7.10^{-5} sec. were obtained. It is of course obvious that the resolving-time depends on the grid time-constant of the pentode as determined by the appropriate resistances and capacities, but in addition it will be seen that the resolving-time will be improved if the input pulse is smaller so that the pentode remains a shorter time in the region of zero current. This result can be achieved by decreasing the over-voltages on the counters but this must not be carried too far, otherwise there will be a noticeable loss of coincidences due to a coincidence occurring while one counter has not yet recovered sufficiently from a previous discharge; see figure 4. Owing to this effect the resolving-power will depend on the number of single pulses in each counter.

τ_c
D
N_1, N_2

§ 6. PULSE-SHAPE

The shape of pulse developed may best be discussed by means of a simple equivalent circuit shown in figure 6. Distortion caused by operating across non-linear regions of the characteristic curve is not considered, because it does not seriously affect the final form of the pulses.

R_i, C_a The pentode is represented by switch D. R_a, R_i and C_a denote external resistance, internal resistance and plate capacity respectively. C is the coupling-

C_t, R_t condenser to the thyratron with grid capacity C_t and grid resistance R_t. As long as there is no pulse, switch D is connected to a. The plate is only slightly above earth-potential, because of the very low internal resistance R_i. The other side of the coupling condenser C is maintained at the grid-bias potential of the thyratron.

The steep wave-front of the discharge pulse is analogous to a sudden opening of switch D across the potentiometer. Plate-capacity C_a plus a small extra capacity (less than $C_t + 1/\omega^2 R_t^2 C_t$, where ω is the fundamental angular frequency of the first

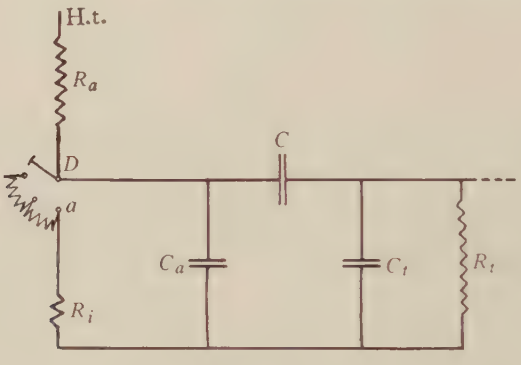


Figure 6.

charge-rise) is quickly charged to the supply voltage, say 150 V. Taking the following approximate values—$C_a \cong 15.10^{-12}$ F., $C_t = 20.10^{-12}$ F., $R_a = 250{,}000\ \Omega$. and $\omega \cong 6.10^5$ (because the incoming pulse rises in a time $t_1 \cong 1.10^{-5}$ sec., § 2), we get for the second term of the above expression 2·6 cm. Hence the whole capacity, which is charged through R_a at the moment of the first rise, is of the order of $15 + 22.10^{-12}$

t_2 or 37.10^{-12} F. If $R_a = 250{,}000\ \Omega$. the time-constant t_2 is $37.10^{-12} \times 25.10^4$, i.e.

t_1 about 10^{-5} sec. and is somewhat smaller than the time t_1 of the first pulse-rise. This is of importance in this circuit, since if t_2 were greater than t_1 not all the pulses would rise to the full h.t. supply voltage of 150 V., and the limiting effect of the characteristic would become disturbed. The back pulse also, as will be seen later, would be eliminated. The conclusion from this consideration is that R_a should not exceed the value of 250,000 Ω.

This positive charging pulse, which is to set off the thyratron, is however reduced in size by the fact that an e.m.f. with resistance R_a is acting on an impedance, comprising C_a, C, C_t and R_t. This impedance is somewhat smaller than R_a. In a scale-of-two circuit R_t is still further reduced by the grid current of the tube when this is

alight. Actually the amplitude of the pulse on the thyratron grid itself is reduced to about 30 or 40 V., a value which is quite sufficient, however, to overcome a normal bias of a few volts.

After passing the lower knee of the curve, figure 3, the system remains unperturbed for a very short time corresponding to the peak of the input pulse. In the equivalent circuit this is analogous to the time during which the switch is open. While the left side of C is still at high potential, the positively charged thyratron grid side starts to discharge through R_t with time-constant $R_t\,(C+C_t)$ and develops simultaneously a potential difference across C. This goes on until the working-point returns from the zero-current region to the active characteristic part with a time-constant determined by the input constants of the pentode grid. In the equivalent circuit this instant is represented by the connection of D to the highest-resistance point of the R_i potentiometer. The potential E_c that has developed across C now becomes divided proportionally between R_i and R_t and the grid of the thyratron

Figure 7.

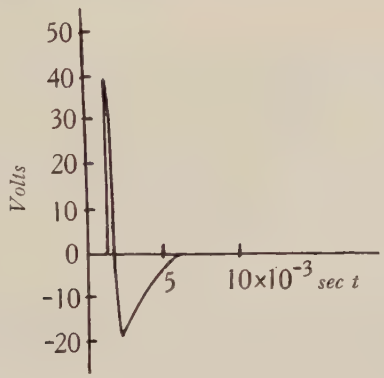


Figure 8.

must receive a negative potential given by $E_c\,R_t/(R_i+R_t)$. The thyratron grid, now negatively charged by the back pulse, discharges with a time constant $R_t\,(C+C_t)$. The pulse finally takes the form shown in figure 7 and represents a very favourable shape for triggering thyratron valves.

In a recent paper[6] Morack has pointed out that this type of pulse is highly desirable. He was able to generate such a pulse by constructing a special thyratron input transformer with iron of several permeabilities. The back pulse helps to accelerate the de-ionization of the excited gas molecules after extinction of the arc. Lewis[7] has emphasized that a sharp, clean pulse can help in preventing the jamming effect. It was found experimentally that with the present amplifier this trouble is completely avoided, even at high rates of counting. It is also to be expected that this pulse-shape, when applied to a thyratron-couple with the improved anode-grid coupling circuit described by Lewis, may lead to a higher degree of resolution for high-speed counting than is at present available.

Oscillograph records were taken of the output pulses from the amplifier which would normally be fed to the thyratron grids. One such record of two close pulses is

shown in figure 7. Owing to the rapid movement of the oscillograph spot the original photographs were not suitable for reproduction and the plate was obtained by careful retouching. A scale drawing of one of the pulses is given in figure 8. The actual rise and return to zero occurring within about 10^{-4} sec. would, if drawn correctly, be indistinguishable on the diagram, but they are shown separated for convenience. This superposition explains why this portion of the trace could just be seen on the oscillograph record whereas the remainder of the downstroke below the axis was invisible. It will be noticed that the back kicks of the two pulses shown in the plate are of different sizes; they were actually 19 V. and 15 V. It is clear from this that the two pulses must have occurred within the recovery time R_0C_0 of the counter, the actual kick from the counter being less for the second pulse, but yet the oscillograph record shows that the initial short positive pulse on the thyratron grids was in both cases the same, about 40 V. This is a good proof of the equalizing action of the amplifier.

§ 7. ACKNOWLEDGMENTS

In conclusion the author wishes to express his thanks to Lord Rutherford for permission to work in the Cavendish Laboratory and to Dr C. D. Ellis for his continued interest and encouragement. He is also greatly indebted to Mr W. J. Henderson for many fruitful discussions and help in the building and testing of the amplifier.

REFERENCES

(1) Johnson and Street. *J. Franklin Inst.* **215**, 239 (1933).
(2) Sven Werner. *Z. Phys.* **92**, 705 (1934), and **90**, 384 (1934).
(3) Pidgeon, H. A. "Theory of Multielectrode Tubes," *Elect. Engng*, Lond., p. 1488 (November 1934).
(4) Rossi. *Nature*, Lond., **125**, 636 (1930).
(5) Fussel and Johnson. *J. Franklin Inst.* **217**, 517 (1934).
(6) Morack, H. M. *Gen. elect. Rev.* **37**, 288 (1934).
(7) Lewis, W. B. *Proc. Camb. phil. Soc.* **30**, 543 (1934).

548·73:546·47-31

THE LATTICE-DIMENSIONS OF ZINC OXIDE

By C. W. BUNN

Received March 16, 1935. *Read in title June* 21, 1935

ABSTRACT. The lattice-dimensions of pure zinc oxide condensed from the smoke have been determined to a high degree of accuracy by means of X-ray powder photographs. Bradley and Jay's lattice-dimensions of quartz[4] and Siegbahn's X-ray wavelengths[5] being accepted as standards, the lattice-dimensions of zinc oxide are found to be:

$$\left.\begin{aligned} a_0 &= 3{\cdot}2426_5 \pm 0{\cdot}0001 \\ c_0 &= 5{\cdot}1948 \pm 0{\cdot}0003 \\ \text{Axial ratio } c_0/a_0 &= 1{\cdot}6020_0 \pm 0{\cdot}0001 \end{aligned}\right\} \text{at } 18^\circ \text{ C.}$$

Both the lattice-dimensions themselves and the axial ratio are lower than Finch and Wilman's electro-diffraction values[2] which are based on the X-ray lattice-dimension of gold, by amounts larger than the combined errors of both methods. There are thus real discrepancies, which are of the order of $\frac{1}{2}$ to 1 per cent.

In view of these discrepancies, Finch and Wilman's assumption that the apparent lattice-dimension of gold is the same for 40–60 kV. electrons as for X-rays may not be correct. The discrepancies therefore represent the combined possible discrepancies shown by gold and zinc oxide.

This does not apply to the axial ratio, which is independent of any reference substance. The difference between the axial ratios determined by electrons and by X-rays appears to point to a real difference in axial ratio between the surface and the interior of zinc oxide crystals.

§ 1. INTRODUCTION

Pure zinc oxide condensed from the smoke has been selected by Finch and Quarrell[1] as a reference substance in electron-diffraction determinations of crystal-lattice dimensions, and Finch and Wilman have recently determined its lattice-dimensions by reference to gold[2] using 40 to 60 kV. electrons. The values obtained in this way are considerably higher than those obtained previously by X-ray methods. The figures are collected in table 1, which includes a set of X-ray results that have appeared since Finch and Wilman's work.

Although the X-ray values differ considerably among themselves, they are consistently lower than the electron-diffraction values. Finch and Wilman conclude that the difference between the X-ray and electron-diffraction figures is due to a difference in the nature of the zinc oxide examined, and state that "this question could probably be settled, and the interests of accurate electron-diffraction analysis materially furthered, by an X-ray examination of pure condensed zinc-oxide smoke." This has now been done.

Table 1

Material	a_0 (A.)	C_0 (A.)	Axial ratio	Reference
X-ray method				
Zincite	3·22	5·20	1·608	W. L. Bragg[9]
Zincite	3·22	5·20	1·608	Aminoff[10]
Zincite	3·251	5·226	1·6077	Weber[11]
Zinc oxide (Merck)	3·242	5·176	1·596	Barth[12]
Zinc oxide (spectroscopically pure)	3·235	5·209	1·610	Fuller[13]
Zinc oxide	3·248	5·202	1·602	Ivannikov, Frost, Shapiro[14]
Electron-diffraction method				
Zinc-oxide smoke	3·258	5·239	1·607	Finch and Wilman[2]

§ 2. EXPERIMENTAL

Zinc-oxide smoke was made by the combustion of pure granulated zinc of forensic quality supplied by Prof. Finch. X-ray powder patterns were obtained by the Hull-Debye-Scherrer method, using a camera 9 cm. in diameter of the type described by Bradley and Jay[3]. The camera was improved by the addition of brass screens which trap the primary beam soon after it has passed the specimen at the centre of the camera, the dimensions being such that none of the scattering which occurs inside the trap can possibly reach any part of the film; among other things, this has the effect of reducing the general background intensity, thus increasing the contrast of the lines and materially aiding the attainment of the highest accuracy of measurement.

In Bradley and Jay's type of camera, the exposed portion of the film is brought to an abrupt end by a knife edge, which thus registers a constant angle on every film. In the present instance this angle corresponds to a Bragg angle of 82°. For accurate determinations of lattice constants, only the reflections at large Bragg angles of 60 to 80°, which are well resolved into α_1 and α_2 doublets, are used. The distances between pairs of corresponding lines and the total length of the exposed portion of the film are measured, the angles being obtained by simple proportion. This eliminates errors due to shrinkage of the film.

The constant camera-angle is found by taking powder photographs of a substance whose lattice-dimensions are accurately known; Bradley and Jay have shown[4] that quartz is by far the best for this purpose. A sample of pure quartz, similar to that used by Bradley and Jay, was kindly supplied by Dr H. E. Buckley. Four powder photographs were taken with copper Kα radiation, four separately mounted specimens being used so that the corrections should be different. The camera angle θ_k was found by calculating a value from each reflection in the range 60 to 80°, plotting these values against $\sin^2 2\theta$ 2θ, and extrapolating to $\sin^2 2\theta = 0$. The mean value of θ_k from the four films was found to be $82{\cdot}125° \pm 0{\cdot}005°$. This error of $\pm$ 1 in 16,000 in the value of θ means that if the above value of θ_k is assumed, the lattice-dimensions of quartz can be obtained by the following method with an error of $\pm$ 1 in 60,000.

The lattice constants of a hexagonal crystal are found by first calculating an approximate value for the axial ratio c from several pairs of reflections by means of the equation

$$(d_{hkl})^2 \{\tfrac{4}{3}(h^2+k^2+hk)+l^2/c^2\} = (d_{h_1k_1l_1})^2 \{\tfrac{4}{3}(h_1^2+k_1^2+h_1k_1)+l_1^2/c^2\}.$$

This axial ratio is then used to calculate a_0 from all the reflections at large angles, thus a_0

$$a_0 = d_{hkl}\sqrt{\left\{\frac{4}{3}(h^2+k^2+hk)+\frac{l^2}{c^2}\right\}}.$$

These values of a_0 are plotted against $\cos^2\theta$, and a linear extrapolation is made to $\cos^2\theta = 0$. Bradley and Jay have shown that errors due to absorption of X-rays, thickness of specimen, and eccentricity of mounting are eliminated in this way. If there is any systematic one-sided deviation from the line on the part of values derived from reflections with large l indices, the axial ratio is in error, and another value must be tried.

In the present instance two powder photographs were taken with copper Kα radiation and two with cobalt Kα radiation, all with separately mounted specimens so that the corrections should be different. The zinc-oxide smoke gave very good photographs, with strong sharp lines.

There is no sign of line-broadening, and the crystals are therefore larger than 1000 A. in diameter. It so happens that with both radiations there is one strong doublet very near the end of the film at an angle of over 80°, and there are several more between 60° and 80°; this makes for great accuracy.

The films were measured on a Pye travelling microscope reading to 0·001 cm. Values of d/n were calculated from the Bragg equation $d/n = \lambda/2 \sin\theta$ for each line. The pairs of values from α_1 and α_2 wave-lengths were averaged, the α_1 value being given twice the weight of the α_2 value on account of its greater intensity. A correction for refraction of X-rays was then made by means of Siegbahn's formula[5]:

$$\text{True } d = d \text{ (observed)} \times \left(1 + 5{\cdot}4\,\frac{d^2\rho}{n^2}\cdot 10^{-6}\right),$$

where ρ is the density of the crystal (5·68 for ZnO). ρ

The indices of the reflections were found in the usual way, by means of a chart of the Hull and Davey type. The lattice constants were then found in the way already described. Details are given in tables 2 and 3. The value of the axial ratio c which best fitted all the films was found to be $1{\cdot}6020_0$. The error here is not likely to be greater than $\pm 0{\cdot}0001$.

The temperature near the camera was about 18° C. for all the films.

It may be noted that the final cobalt values are higher than the final copper values. The cobalt values could be reduced to the same level as the copper values by adopting a slightly higher axial ratio (1·6021), since the value from 213, which chiefly determines the final cobalt value, would be reduced by about the right amount; but a higher axial ratio would throw some of the other values too far out, and therefore has not been adopted.

Table 2. Powder photographs of ZnO taken with copper Kα radiation

$\alpha_1 = 1{\cdot}537395$ A. $\alpha_2 = 1{\cdot}541232$ A.

Reflection	Intensity	Film I					Film II				
		θ (deg.)	$\frac{d}{n}$ (A.)	$\frac{d}{n}$ average	$\frac{d}{n}$ corrected	a_0	θ (deg.)	$\frac{d}{n}$ (A.)	$\frac{d}{n}$ average	$\frac{d}{n}$ corrected	a_0
213 α_1	Very strong	58·247	0·90401	0·90405	0·90408	3·2396	58·226	0·90421	0·90419	0·90422	3·2401
α_2		58·464	0·90415				58·463	0·90416			
302 α_1	Strong	60·884	0·87988	0·87987	0·87989	3·2400	60·880	0·87992	0·87998	0·88000	3·2403
α_2		61·145	0·87986				61·117	0·88009			
006 α_1	Weak	62·676	0·86524	0·86524	0·86526	3·2407	—	—	—	—	—
205 α_1	Strong	67·075	$0{\cdot}834626$	$0{\cdot}834657$	$0{\cdot}834672$	$3{\cdot}24070$	67·211	$0{\cdot}834855$	$0{\cdot}834859$	$0{\cdot}834876$	$3{\cdot}24149$
α_2		67·398	$0{\cdot}834726$				67·375	$0{\cdot}834861$			
106 α_1	Medium weak	68·372	$0{\cdot}826918$	$0{\cdot}826918$	$0{\cdot}826932$	$3{\cdot}24098$	68·364	$0{\cdot}826968$	$0{\cdot}827001$	$0{\cdot}827016$	$3{\cdot}24131$
α_2		—	—	—	—	—	68·712	$0{\cdot}827075$			
214 α_1	Medium	69·355	$0{\cdot}821450$	$0{\cdot}821440$	$0{\cdot}821456$	$3{\cdot}24111$	69·350	$0{\cdot}821476$	$0{\cdot}821490$	$0{\cdot}821514$	$3{\cdot}24137$
α_2		69·743	$0{\cdot}821421$				69·720	$0{\cdot}821541$			
220 α_1	Strong	71·552	$0{\cdot}810341$	$0{\cdot}810361$	$0{\cdot}810370$	$3{\cdot}24152$	71·547	$0{\cdot}810361$	$0{\cdot}810403$	$0{\cdot}810418$	$3{\cdot}24167$
α_2		71·971	$0{\cdot}810407$				71·955	$0{\cdot}810481$			
310 α_1	Strong	80·776	$0{\cdot}778770$	$0{\cdot}778766$	$0{\cdot}778779$	$3{\cdot}24231$	80·761	$0{\cdot}778802$	$0{\cdot}778780$	$0{\cdot}778793$	$3{\cdot}24237$
α_2		81·711	$0{\cdot}778752$				81·728	$0{\cdot}778717$			
Extrapolated a_0						$3{\cdot}24262$					$3{\cdot}24260$

Table 3. Powder photographs of ZnO taken with cobalt Kα radiation

$\alpha_1 = 1{\cdot}78529$ A. $\alpha_2 = 1{\cdot}78919$ A.

Reflection	Intensity	Film III					Film IV				
		θ (deg.)	$\frac{d}{n}$ (A.)	$\frac{d}{n}$ average	$\frac{d}{n}$ corrected	a_0	θ (deg.)	$\frac{d}{n}$ (A.)	$\frac{d}{n}$ average	$\frac{d}{n}$ corrected	a_0
211 α_1	Strong	59·269	1·03847	1·03852	1·03855	3·2384	59·227	1·03893	1·03892	1·03895	3·2396
α_2		59·468	1·03860				59·440	1·03890			
114 α_1	Medium strong	61·843	1·01246	1·01255	1·01258	3·2394	61·817	1·01271	1·01280	1·01283	3·2402
α_2		62·051	1·01271				62·023	1·01297			
212 α_1	Medium strong	65·409	$0{\cdot}98168_2$	$0{\cdot}981714$	$0{\cdot}98174_3$	$3{\cdot}2400_4$	65·375	$0{\cdot}98195_0$	$0{\cdot}98199_7$	$0{\cdot}98202_6$	$3{\cdot}2409_7$
α_2		65·672	$0{\cdot}98177_5$				65·631	$0{\cdot}98209_4$			
105 α_1	Strong	66·461	$0{\cdot}97366_7$	$0{\cdot}97371_3$	$0{\cdot}97374_2$	$3{\cdot}2404_6$	66·425	$0{\cdot}97393_3$	$0{\cdot}97392_6$	$0{\cdot}97395_5$	$3{\cdot}2411_7$
α_2		66·731	$0{\cdot}97380_5$				66·716	$0{\cdot}97391_5$			
204 α_1	Weak	69·544	0·95272	0·95272	0·95275	3·2404	69·496	0·95302	0·95302	0·95305	3·2415
300 α_1	Strong	72·564	$0{\cdot}93563_6$	$0{\cdot}93569_2$	$0{\cdot}93571_9$	$3{\cdot}2414_3$	72·535	$0{\cdot}93578_5$	$0{\cdot}93581_3$	$0{\cdot}93583_9$	$3{\cdot}2418_4$
α_2		72·934	$0{\cdot}93580_2$				72·921	$0{\cdot}93586_7$			
213 α_1	Very very strong	80·592	$0{\cdot}90481_6$	$0{\cdot}90483_4$	$0{\cdot}90485_7$	$3{\cdot}2424_0$	80·577	$0{\cdot}90485_5$	$0{\cdot}90486_0$	$0{\cdot}90488_2$	$3{\cdot}2424_8$
α_2		81·358	$0{\cdot}90487_0$				81·358	$0{\cdot}90487_0$			
Extrapolated a_0						$3{\cdot}2427_4$					$3{\cdot}2427_2$

The discrepancy is probably due to refraction effects which cannot be precisely evaluated. Siegbahn's correction formula is for symmetrical reflection, i.e. the incoming and outgoing rays make equal angles with the crystal surface. This condition is usually not fulfilled on account of the development of a limited number of faces on the crystals; each line in a powder photograph is made up of many individual reflections which undergo varying deviations, the resultant of which may not correspond with that given by the formula (6). The last line on the cobalt films is from a totally different plane from that on the copper films, and the refraction corrections may not be quite right for either of them. However, since the extrapolated values of a_0 lie within the expected limits of error, this question need not be considered further; any inaccuracies due to these refraction effects are not likely to throw the final average value outside the limits given.

In the films taken with copper radiation, the last two reflections are from prism planes (310 and 220), and hence the extrapolated value of a_0 depends very little on the axial ratio adopted. These values have therefore been allowed double weight in taking a final average of a_0, which is found to be as follows:

$$a_0 = 3{\cdot}2426_5 \pm 0{\cdot}0001 \text{ A.}$$

Combining this with the axial ratio $1{\cdot}6020 \pm 0{\cdot}0001$, we get

$$c_0 = a_0 c = 5{\cdot}1947 \pm 0{\cdot}0004 \text{ A.}$$

In giving a possible error of ± 1 in 30,000 in the value of a_0, Bradley and Jay's lattice-dimensions of quartz are taken as standards. These in turn rest on Siegbahn's X-ray wave-lengths. Any error here will, of course, increase the absolute error of the present results. But for the present purpose of comparison with Finch and Wilman's electron-diffraction results, this does not enter into the problem, since those workers used as standard the lattice-dimension of gold, which rests on the same X-ray wave-lengths.

The c_0 dimension can be obtained with a little more precision by following the same procedure as for a_0, by means of the equation

$$c_0 = d\sqrt{\{\tfrac{4}{3}(h^2+k^2+hk)\,c^2+l^2\}}.$$

The results are given in table 4.

Table 4. Powder photographs of ZnO. Determination of c_0

Reflection	c_0 (A.)		Reflection	c_0 (A.)	
	Film I	Film II		Film III	Film IV
213	$5{\cdot}1898$	$5{\cdot}1906$	—	—	—
302	$5{\cdot}1904$	$5{\cdot}1910$	211	$5{\cdot}1879$	$5{\cdot}1899$
006	$5{\cdot}1915$	—	114	$5{\cdot}1895$	$5{\cdot}1908$
205	$5{\cdot}1916$	$5{\cdot}1928_8$	212	$5{\cdot}1905_5$	$5{\cdot}1920_4$
106	$5{\cdot}1920_5$	$5{\cdot}1925_8$	105	$5{\cdot}1912_2$	$5{\cdot}1923_6$
214	$5{\cdot}1923_0$	$5{\cdot}1926_7$	204	$5{\cdot}1911_8$	$5{\cdot}1928_0$
220	$5{\cdot}1929_1$	$5{\cdot}1931_6$	300	$5{\cdot}1927_7$	$5{\cdot}1934_4$
310	$5{\cdot}1941_9$	$5{\cdot}1942_9$	213	$5{\cdot}1943_1$	$5{\cdot}1944_6$
	$5{\cdot}1946_5$	$5{\cdot}1946_8$		$5{\cdot}1949_2$	$5{\cdot}1948_6$

The results from films III and IV, taken with cobalt radiation, depend on the axial ratio to a smaller extent than the other results. Allowing them therefore double weight in taking a final average, we obtain

$$c_0 = 5{\cdot}1948 \pm 0{\cdot}0003 \text{ A.}$$

The limits of error for c_0 are wider apart than for a_0; this is because the lines nearest to the ends of the films are not basal-plane reflections.

§ 3. DISCUSSION

The lattice-dimensions found by the X-ray method differ from those found by the electron-diffraction method for similar material by amounts larger than the combined errors of both methods.

Table 5

	a_0	c_0	Axial ratio
X-ray method	$3{\cdot}2426_5 \pm 0{\cdot}0001$	$5{\cdot}1948 \pm 0{\cdot}0003$	$1{\cdot}6020_0 \pm 0{\cdot}0001$
Electron-diffraction method	$3{\cdot}258 \pm 0{\cdot}005$	$5{\cdot}239 \pm 0{\cdot}005$	$1{\cdot}607 \pm 0{\cdot}002$

There is thus a real difference between X-ray and electron-diffraction values; and not only are the lattice-dimensions themselves different, but also the axial ratios are different.

It must be remembered, however, that the electron-diffraction values for the lattice-dimensions rest on the assumption that the electron-diffraction lattice-dimension of gold is the same as the X-ray value. It now appears that this assumption may not be correct; if zinc oxide shows a discrepancy, gold may do the same. The discrepancies found here represent the combined differences of gold and zinc oxide. This does not apply to the axial ratio, which is independent of the reference substance.

The discrepancies recall the much larger discrepancies (up to 30 per cent) found by Davisson and Germer[7] who used slow electrons with metallic crystals. On that occasion the apparent lattice-contraction was at first thought to indicate a real contraction on the crystal surface, which is revealed by electron waves on account of their very small penetration. But according to the calculations of Lennard-Jones and Dent[8] the surface contraction on ionic crystals of the NaCl type is only of the order of 5 per cent, and is practically confined to the first layer of atoms; if the same applies to metal crystals, this explanation seems unlikely to be correct. It was considered that a more likely explanation of the apparent lattice-contraction was to be found in the refraction of electron waves by the crystal.

Here we have to account for apparent expansions of the order of $\frac{1}{2}$ to 1 per cent when fast (40 to 60-kV.) electrons are used. It is not very profitable to carry the discussion further at present, for there is insufficient evidence to point to a definite conclusion. Even the sign of the change in ZnO is uncertain, since the behaviour of gold is unknown. One fact, however—the difference between the axial ratios

measured by the two methods—does appear to point to a difference in the crystalline material encountered by X-rays and by electrons. Since the penetrating power even of fast electrons is very much less than that of X-rays, this indicates a change of axial ratio at the surface. Any surface change of lattice dimensions would, of course, affect the *a* and *c* axes in different ratios.

§ 4. ACKNOWLEDGMENT

The author wishes to express his thanks to the Directors of Imperial Chemical Industries Limited for permission to publish this work, which was carried out in the Research Laboratories of their subsidiary company I.C.I. (Alkali) Limited.

REFERENCES

(1) Finch and Quarrell. *Proc. phys. Soc.* **46**, 148 (1934).
(2) Finch and Wilman. *J. chem. Soc.* p. 751 (1934).
(3) Bradley and Jay. *Proc. phys. Soc.* **44**, 563 (1932).
(4) Bradley and Jay. *Proc. phys. Soc.* **45**, 507 (1933).
(5) Siegbahn. *Spektroskopie der Röntgenstrahlen*, Eng. Trans. 1925, p. 26.
(6) Hägg and Phragmen. *Z. Krystallogr.* **86**, 306 (1933).
(7) Davisson and Germer. *Nature*, Lond., **119**, 558 (1927).
(8) Lennard-Jones and Dent. *Proc. roy. Soc.* **121**, 247 (1928).
(9) W. L. Bragg. *Phil. Mag.* **39**, 647 (1920).
(10) Aminoff. *Z. Krystallogr.* **56**, 495 (1921); **57**, 204 (1922).
(11) Weber. *Z. Krystallogr.* **57**, 398 (1922).
(12) Barth. *Norsk geol. Tidsskr.* **9**, 317 (1927).
(13) Fuller. *Science*, **70**, 196 (1929).
(14) Ivannikov, Frost and Shapiro. *C.R. Acad. Sci. U.S.S.R.* p. 124 (1933).

551·51·053·5

THE STRUCTURE OF THE IONOSPHERE

By Professor J. HOLLINGWORTH, M.A., D.Sc., M.I.E.E.

Received December 18, 1934. *Read in title June* 21, 1935

ABSTRACT. The following paper suggests a method of giving an approximate value to the intensity of ionization in the space between the E and F layers by measuring the relative group retardation of the two components of an echo from a pulse whose frequency is so close to the critical frequency of the E region that one component is reflected from the E layer and one from the F. It is pointed out that the value of the inter-layer ionization is of the utmost importance in determining the path of long-distance transmissions. From the results it is concluded that probably this inter-layer ionization is fairly uniformly distributed over the space between the layers, and is only a few per cent less in value than that at the top of the E layer.

§ 1. INTRODUCTION

As a result of the experiments of Appleton and others which have resulted in the magneto-ionic theory of wireless wave-propagation, it was originally held that the ionization of the atmosphere broadly consisted of two layers: the E or Kennelly-Heaviside layer at a height of about 100 km. above the surface of the earth, and the F or Appleton layer about 150 km. higher. The maximum ionization in the E layer being less than that in the F, signals for which the ionic content of E is not sufficient to return them break through abruptly to F, and this was verified by both the wave-length change and the pulse methods.

The actual value of the ionization in this intermediate region is, however, a decisive factor in the propagation of long-distance signals, as will be seen from the following argument. A signal of sufficiently high frequency to penetrate the E layer will, of course, be deflected by this layer; but if it emerges from it into an un-ionized region it will again revert to its original direction until it strikes the F layer, whereas if this intermediate space be ionized its direction at any point will be determined by the appropriate value of μ at that point.

The earlier workers on long-distance propagation adopted the original idea of an un-ionized space, since in their diagrams they show the signal path to the F layer as triangular, possibly with the apex slightly rounded off. On this supposition one would expect the principal ray to strike the earth at nearly grazing incidence; the exact angle being determined by a geometric triangulation between transmitter and receiver. Early in 1931, however, the author was engaged in measuring the angles of incidence of the signals from certain long-distance stations, and the results indicated that these angles were smaller, sometimes very much so, than this theory would lead one to expect. As a possible explanation it was suggested that the space between the E and F layers may be ionized to a degree only slightly less than that at the top of the E layer. Under these conditions the principal ray would be that which, subject to the necessary geometrical limitations, emerges from the

top of the *E* layer at nearly grazing incidence, and consequently leaves the earth more steeply. If this were the case, the lower the frequency the more pronounced would be the effect, subject of course to the condition that the frequency was high enough to penetrate the *E* layer completely. On the waves of 15 to 20 metres used for commercial purposes the *E*-layer deflection could only be expected to be a few degrees, so that the effect would be much less marked and might easily be overlooked.

Recent measurements appear to have confirmed this supposition by showing angles of incidence of 75 to 85 degrees[2, 3].

The same idea has since been put forward in various papers by other workers, chiefly Appleton and Ratcliffe, and is now generally accepted in principle. Their reasoning was based on the occasional appearance of echoes from a height lying between the accepted heights of the normal *E* and *F* layers; but from the nature of the method employed these echoes can be obtained only when the ionization of this region exceeds that of the top of the *E* layer. If such echoes are not present, all that can be said is that the ionization of this intermediate region is less than that of the top of the *E* layer, but no estimate can be given as to how much less. One investigator specifically refers to the "absence of the ϵ region" on such occasions, and another shows a diagram with this portion of the curve relating ionization with height dotted. For long-distance propagation, however, it has been explained that it is most important to know something of the ionization of this region even when it is not capable of returning echoes, for the value of μ changes very rapidly for densities near the critical one, and in a recent paper by Appleton and Builder[4] an account is given of an effect observed occasionally, which seemed to the author to suggest a possibility of obtaining further information on this point.

In this paper Appleton and Builder show that when pulse observations at a definite frequency are being made on the *E* layer at a time when its ionization is steadily diminishing through the critical value for that frequency, reflections of the ordinary and extraordinary rays which have only just succeeded in penetrating the *E* layer may for a short time be actually reversed in time. They rightly attribute this to the reduction in group-velocity of the extraordinary ray, which can only just penetrate the *E* layer, relative to that of the ordinary ray whose critical density is much higher; and it occurred to the author that it might be possible to make an approximate computation of this time lag and see whether it could be accounted for by a discrete layer only 15–20 km. thick. The results appear to indicate very clearly that this is not possible; that in a layer whose ionization increases from zero at the bottom to the critical value some 15 km. higher up, and then falls off again, the distance over which there is appreciable reduction of group-velocity is so short that the time lag obtained is not comparable with that actually observed.

It must therefore be assumed that the ionization at the top of the *E* layer persists with only a slight diminution in value until the *F* layer is reached.

§ 2. ANALYTICAL METHOD

The standard formula for group-velocity U is U

$$\frac{1}{U}=\frac{1}{c}\frac{d}{dp}(\mu p),$$

c being the velocity of light *in vacuo*, p the periodicity, and μ the refractive index of the medium. c, p, μ

Hence if U and U' are the group-velocities of the ordinary and extraordinary rays, their coefficients of refraction being μ and μ' respectively, we have U' μ'

$$\frac{1}{U}=\frac{1}{c}\frac{d}{dp}(\mu p), \qquad \frac{1}{U'}=\frac{1}{c}\frac{d}{dp}(\mu' p).$$

Now if a velocity U operates over a distance ds, the time of passage is ds/U; i.e. $1/U$ represents the time occupied in travelling unit distance through a medium of constant electron-density of the appropriate value. The time lag between the two components thus becomes $1/U-1/U'$ per unit distance. ds

Therefore the differential time-lag rate between the ordinary and extraordinary rays is given by

$$\frac{1}{c}\left\{\frac{d}{dp}(\mu p)-\frac{d}{dp}(\mu' p)\right\}.$$

If $c=3.10^{10}$ this lag is in seconds per centimetre of path; but a more useful unit for this work is milliseconds per kilometre, in which case we get

Differential delay (msec./km.)

$$\frac{1}{300}\left\{\frac{d}{dp}(\mu p)-\frac{d}{dp}(\mu' p)\right\} \qquad \text{.....(1).}$$

A formula relating μ, p and N, the ionization-density, is given in the paper referred to, and is used here except that in view of recent mathematical investigations the expression $(-p^2/p_0^2-\frac{1}{3})$ for α has been replaced by $-p^2/p_0^2$. N

Subject to this condition the formula reduces to

$$\mu^2=1-\frac{2p_0^2}{2p^2-p^2p_T^2/(p^2-p_0^2)\mp\sqrt{\{p^4p_T^4/(p^2-p_0^2)^2+4p^2p_L^2\}}} \qquad \text{.....(2),}$$

p_T, p_L

where the symbols are as defined in the paper.[4]

There is no fundamental difficulty in evaluating the retardation formula from this expression, but the arithmetical work would be extremely heavy. Moreover the experimental measurement of the retardation times does not admit of a very high order of accuracy, so that we are concerned rather with the order of certain quantities than with their exact values. It is therefore permissible to use approximations throughout, and these have been made on the following basis.

If a rough calculation be made of the required retardation times, it will be found that the delay time when plotted against ionization-density remains quite small until the critical density is nearly reached, when it rises very rapidly. This is also evident from the curve in figure 1. The most important part of the curve is thus at densities differing from the critical density by first-order amounts, so that

in the expression for $d(\mu^2)/dp$ we can after differentiation insert the critical relation for the extraordinary ray, namely

p_H

$$p'^2 - p_0^2 = pp_H.$$

When this has been done, the expression reduces to

$$\frac{d(\mu^2)}{dp} = \frac{2(2p' - p_H)}{p_0^2(1 + p_L^2/p_H^2)},$$

p' and since $\mu = 0$ when p has its critical value p', we find that for any value of p close
μ' to p' we can write for the corresponding index of refraction μ'

$$\mu' = \left\{\frac{2(2p' - p_H)}{p_0^2(1 + p_L^2/p_H^2)}\right\}^{\frac{1}{2}}(p' - p)^{\frac{1}{2}} = \kappa(p' - p)^{\frac{1}{2}} \qquad \ldots\ldots(3).$$

Applying the standard group-velocity formula to this, we find that, since

$$\frac{d(\mu' p)}{dp} = \mu' + p\frac{d\mu'}{dp},$$

and since μ' is small and $d\mu'/dp$ large, the following result

$$\frac{1}{U'} = \frac{\kappa p'}{2c}(p' - p)^{-\frac{1}{2}}$$

is true for values of p close to p'.

For the purpose of evaluating the total delay through a layer of varying density we require the relation between $1/U$ and N, which can be obtained as follows. At the critical value for the extraordinary ray we have

$$p'^2 - p_0^2 = pp_H,$$

where

$$p_0^2 = 4\pi e N'/m.$$

By differentiating this with regard to p we obtain dN in terms of dp, i.e. we get
N' $N' - N$ in terms of $p' - p$ where $N' - N$ is small; so that on substituting in equation (3) we finally obtain the following relation for the group-retardation rate for the extraordinary ray:

A

$$\frac{1}{U'} = \frac{\sqrt{2p'(2p' - p_H)}}{cp_0(1 + p_L^2 p_H^2)^{\frac{1}{2}}A}(N' - N)^{-\frac{1}{2}} \text{ where } A^2 = \frac{4p_0^2}{N} = \frac{16\pi e^2}{m} \qquad \ldots\ldots(4).$$

Actually we require the difference between this and the rate for the ordinary ray; but as the latter is far from its critical frequency the approximate formula cannot be used. Another approximation is however possible here. In the original formula for μ it is found that when $p^2 > p_0^2$ the term $p^4 p_T^4 (p^2 - p_0^2)^2$ rapidly becomes small compared with $4p^2 p_L^2$. On this assumption the formula reduces to

$$\mu^2 = 1 - \frac{p_0^2}{p^2 + pp_L},$$

and by a process similar to the previous one we get for the delay rate of the ordinary ray

$$\frac{1}{U} = \left(2p - \frac{p_0^2 p_L}{p^2 + p_L^2}\right) \Big/ p.c.\left(1 + \frac{p_0^2}{p^2 + p_L^2}\right)^{\frac{1}{2}} \qquad \ldots\ldots(5).$$

There will, of course, be an intermediate stage in which neither of these approximations is valid. Here it is best to calculate a few points as accurately as possible and to use them to correct the other curves; but the whole of this range lies in a region where the delay is small, so that the possible error will not have any serious effect.

§ 3. NUMERICAL EVALUATION

In these calculations we shall assume that $p = 22 \times 10^6$ radians/sec., giving a wave-length of 85·7 metres which is of the same order as that used by Appleton and Builder. The critical density for the ordinary ray is then $1{\cdot}51 \times 10^5$, and for the extraordinary $0{\cdot}9467 \times 10^5$ electrons/cm.^{3}

Other constants for vertical transmission in the latitude of Slough are as follows:

$$p_H = 8{\cdot}302 \times 10^6, \qquad p_L = 7{\cdot}643 \times 10^6, \qquad p_T = 3{\cdot}24 \times 10^6.$$

Performing the necessary numerical work we finally arrive at the following results:

For values of N near to the critical value N', for the extraordinary ray,

$$\mu' = 0{\cdot}00338\,(N' - N)^{\frac{1}{2}} \qquad \text{.....(6),}$$

$$1/U' = 1{\cdot}383/(N' - N)^{-\frac{1}{2}} \text{ (msec./km.)} \qquad \text{.....(7).}$$

For values of N far from N',

$$\mu = 1 - 0{\cdot}488\,.N.10^{-5} \text{ (ordinary ray)} \qquad \text{.....(8),}$$

$$\mu' = 1 - 1{\cdot}007\,.N.10^{-5} \text{ (extraordinary ray)} \qquad \text{.....(9),}$$

$$\frac{1}{U} = \frac{44 \times 10^6 - 27{\cdot}8\,.N}{(48{\cdot}4 \times 10^{13} - 23{\cdot}2 \times N \times 10^8)^{\frac{1}{2}}} \cdot \frac{1}{2c} \text{ (ordinary ray)} \qquad \text{.....(10),}$$

$$\frac{1}{U'} = \frac{44 \times 10^6 + 118\,.N}{(48{\cdot}4 \times 10^{13} - 48{\cdot}7 \times N \times 10^8)^{\frac{1}{2}}} \cdot \frac{1}{2c} \text{ (extraordinary ray)} \qquad \text{.....(11).}$$

The numerical results are given in the accompanying tables.

Table 1 (Ordinary ray)

$N \times 10^{-5}$	0·2	0·4	0·6	0·8	0·9	1·0
μ, equation (8)	0·95	0·9	0·84	0·78	0·75	0·72
μ, equation (2)	—	—	—	0·776	—	0·743
$1/U$	0·00347	0·0036	0·00383	0·00403	0·00413	0·00427

In order to check the accuracy of these approximate formulae the values of μ and $1/U$ have been tabulated so that the results given by the formulae overlap; and also a few values of μ calculated from the fundamental formula have been added.

Referring first to the ordinary ray, table 1, it will be seen that over the range considered the approximate and absolute values of μ agree so closely that there is little reason to doubt the accuracy of the $1/U$ formula; and in any case the delay on this ray is so small that slight errors have a negligible over-all result.

For the extraordinary ray, table 2, the calculation had to be made for a much larger number of values of N, owing to the approach to the critical point; but in the overlap period ($N = 0{\cdot}8 \times 10^5$ and $0{\cdot}9 \times 10^5$) comparison with the accurate figures suggests that those results which are enclosed in brackets should be rejected.

As regards the relative values for $N = 0{\cdot}94 \times 10^5$ it may be noted that it is quite likely that here the result from the approximate formula is more accurate than that from the general one, for when $N' - N$ is very small the expression for μ^2 in the general formula involves the small difference between two large numbers, and so requires an extremely large number of significant figures to yield high accuracy.

Table 2 (Extraordinary ray)

$N \times 10^{-5}$	0·2	0·4	0·6	0·8	0·9	0·94
μ, equation (9)	0·89	0·77	0·62	[0·44]	[0·3]	—
μ, equation (2)	—	—	—	0·4	0·23	0·905
μ, equation (6)	—	—	—	0·433	0·23	0·874
$1/U$, equation (10)	0·00392	0·00477	0·0058	[0·0092]	0·0134	—
$1/U$, equation (7)	—	—	—	0·0014	0·0203	0·0503

The final results for $1/U - 1/U'$ are plotted in figure 1, and give the differential group-retardation rate between the rays for the given range of densities. Now the

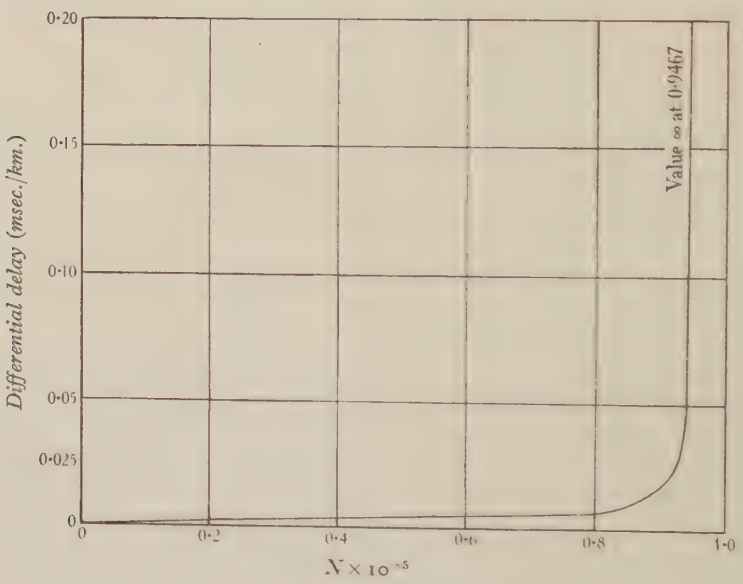


Figure 1.

exact distribution of density with height in the layer is not known, and it will be assumed that the density is proportional to the height above the point at which ionization commences, since this will at any rate give an approximation to the real

figures. Below this height, of course, the differential retardation is always zero. Now if we have a layer graded in this way with a maximum density N', the total delay will be twice the integral of this curve up to the point $N=N'$ multiplied by the thickness of the layer, the factor 2 arising from the fact that the ray passes twice through the layer. This integration can be performed graphically to within a short distance of the critical density, where the curve becomes too steep for accurate evaluation.

It is, however, important to note that although the delay rate is infinite at the critical density yet in a layer graded up to this density the total delay time is finite, since it is an integral of the form $\int \frac{dN}{(N-N')^{\frac{1}{2}}}$, so that the total delay time can be calculated right up to this value, and the results of this calculation are given in

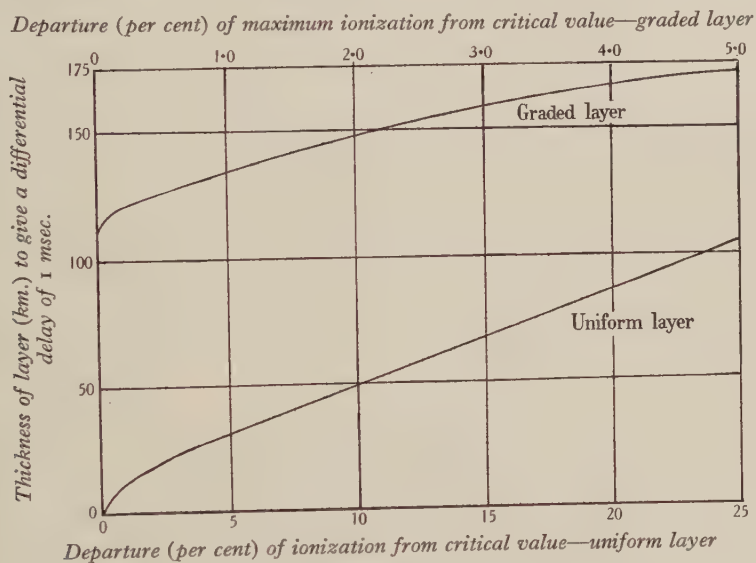


Figure 2.

the third column of table 3. For convenience the densities are given not in their absolute values but as percentage departures from the critical density, $N=0{\cdot}9467\times 10^5$, since in this form they are probably roughly applicable to other critical densities of the same order. An inspection of the traces given by Appleton and Builder will show that the delays under consideration are of the order of 1 msec.

To get the above results in the most suitable form, the figures in table 3 have been used to determine the thickness of layer required, both uniform and graded, to give a delay of 1 msec., and the figures obtained have been plotted in figure 2.

On examination of these results it will be seen that to obtain the necessary delay with a graded layer, even if the maximum value is *exactly* the critical value, would require a thickness of 106 km. Now from the fact that signals returned from the *E* layer show only small differences in their delay times over a considerable range of frequencies, it is clear that the graded part of this layer must be com-

paratively thin. Consequently to obtain a delay of the observed value the ray must pass through a much greater distance of ionized air.

Table 3*

Density (per cent departure from critical value)	Delay (msec. km.)		Thickness giving delay of 1 msec. (km.)	
	Uniform layer	Graded layer	Uniform layer	Graded layer
0	∞	0·00471	0	106·0
0·01 of 1 per cent	0·43	0·0045	1·16	112·0
0·02 ,,	0·308	0·0041	1·63	113·0
0·04 ,,	0·213	0·00436	2·35	114·0
0·07 ,,	0·161	0·00431	3·10	116·0
0·12 ,,	0·122	0·00424	4·10	117·0
0·17 ,,	0·10	0·00418	5·0	120·0
0·50 ,,	0·0591	0·00394	8·46	126·0
0·71 ,,	0·049	0·00383	10·0	130·0
1·23 ,,	0·0366	0·00363	13·7	137·0
1·76 ,,	0·0297	0·00347	17·0	145·0
2·80 ,,	0·0227	0·00322	22·0	155·0
3·87 ,,	0·0187	0·00303	27·0	165·0
4·93 ,,	0·0162	0·00287	31·0	170·0
10·20 ,,	0·0102	0·00227	49·0	220·0
15·40 ,,	0·007	0·00178	71·5	280·0
20·80 ,,	0·006	0·00157	83·0	318·0
26·0 ,,	0·0048	0·00125	109·0	400·0

* All the above results are based on a periodicity of 22×10^6 and a critical frequency for the extraordinary ray of $0{\cdot}9467 \times 10^6$. Probably however they are approximately true over an appreciable range of values.

Moreover the curves show that the delay decreases very rapidly for small departures from the critical density and then changes much more slowly, so that the effect could not be obtained from small patches of layer nearly at the critical density unless we are prepared to admit that the ionization is sufficiently steady for such small patches to retain their ionization for considerable periods of time at a value not differing from the critical value by more than very small fractions of 1 per cent.

The natural assumption appears to be that the space between the *E* and *F* layers is ionized fairly uniformly to a value not much less than that at the top of the *E* layer, or alternatively that the ionization of the layer as a whole rises very rapidly at first for a depth of from 10 to 20 km. and thereafter increases at a much slower rate. The observed delays could then be obtained whenever the density of the layer lies within from 10 to 15 per cent of the critical value, which seems a much more reasonable assumption.

It may be argued that the phenomenon referred to by Appleton and Builder appears to be only rarely observed, and so may be due to an unusual and abnormal state of the layer. But in many of the routine traces taken at the Radio Research Station at Slough it will be noticed that immediately after the break-through to the *F* layer the trace is convex to the origin for a short time, showing an abnormal delay, which is clearly due to the same cause, around the critical frequency. The values of these delays as shown on the traces are generally less than 1 msec.; but owing to the very rapid variation of the delay curve at very small departures from the critical density, extremely accurate instrumental correlation would be needed to obtain the full range of this delay bend. In the apparatus at present in use the wave-length is steadily changing during the emissions, and as the pulses are sent out at the rate of 50 per second not more than one at the utmost would be likely to lie within this highly critical period. For this reason also it seems unprofitable to attempt to build up any exact form of layer-structure from existing observations, though this could probably be done from the calculations given in the present paper if the instrumental survey around the critical frequency could be made of a more rigid nature.

§ 4. ACKNOWLEDGMENTS

In conclusion the author wishes to express his thanks to the Radio Research Board of the Department of Scientific and Industrial Research for permission to make use of some of the results obtained at the Radio Research Station, Slough; and also to Dr Mary Slow and Mr Naismith for their help and suggestions.

REFERENCES

(1) Hollingworth. *J. Instn elect. Engrs*, **72**, No. 435 (March 1933).
(2) Walmsley. *J. Instn elect. Engrs*, **74**, 543 (1934).
(3) Wilkins. *J. Instn elect. Engrs*, **74**, 582 (1934).
(4) Appleton and Builder. *Proc. phys. Soc.* **45**, Part 2, No. 247 (1933).

537.322.2 : 546.74

THE LONGITUDINAL THERMOELECTRIC EFFECT: (II) NICKEL IN LONGITUDINAL MAGNETIC FIELDS

By T. H. PI, M.S. (Yenching) and WILLIAM BAND, M.Sc.,
Professor of Physics, Yenching University, Peiping, China

Received October 1, 1934. *Read in title May* 3, 1935

ABSTRACT. The e.m.f. produced by asymmetrical temperature-distributions in a pure nickel wire is examined and found to agree well with the formula

$$\mathbf{F} = t\mathbf{G} + a\,(\mathbf{G}.\mathbf{G})^{\frac{1}{2}}\,\mathbf{G} + b\,(\mathbf{G}.\mathbf{G})\,\mathbf{G},$$

where **F** is the potential-gradient at the point where **G** is the temperature-gradient. The two constants have the values

$$10^{9}.a = -1{\cdot}80 \text{ V.-cm./(° C.)}^2,$$

$$10^{11}.b = 1{\cdot}85 \text{ V.-cm.}^2/\text{(° C.)}^3,$$

for zero magnetic field. For small fields both constants are reduced in absolute value, reversing sign at 40 gauss, reaching a maximum of this opposite sign at 200 gauss, and thereafter returning asymptotically towards zero. The connexion of these results with thermomagnetic phenomena is discussed.

§ 1. INTRODUCTION

According to the results published in the first paper of this series, on copper[1], the thermoelectric e.m.f. E in a homogeneous wire is expressible simply as a function of the temperature-distribution:

$$E = a\Sigma\,(G.dT) + b\Sigma\,(G^2.dT),$$

where G is the temperature-gradient and T the temperature, and where in the first term on the right-hand side the positive magnitude of G must be used without regard to the sign of the gradient. The potential-gradient thus contains two terms, a quadratic and cubic in the temperature-gradient, in addition to the linear Thomson term.

In the present work we find that the thermoelectric e.m.f. in pure nickel can be represented by the same function with different values of the constants a and b, and that these constants become functions of the magnetizing field when the whole temperature-distribution is placed in a uniform field.

The e.m.f. for copper is opposite in direction to that for nickel, in zero magnetic field, and the values for nickel are much greater than those for copper.

§ 2. APPARATUS AND TECHNIQUE

The specimen was pure thermo-element nickel wire of diameter 0·5 mm. It was stretched by a weight of 500 gm. at each end of the temperature-control system, the specimen making a right-angle bend in two clamps before being taken to the bath, which comprised insulating oil kept at constant uniform temperature.

The temperature-control system consisted of a brass rod with a central bore, at both ends of which were water cooling-jackets at the same cool temperature. The heater was situated close to the inner end of one of the cooling-jackets; it was a nichrome wire wound non-inductively on a truncated slate pyramid surrounding the brass rod coaxially.

This system was adapted to slide into the core of a water-cooled solenoid of the Moullevigen form to give as nearly uniform a field as possible throughout the length of the temperature-distribution.

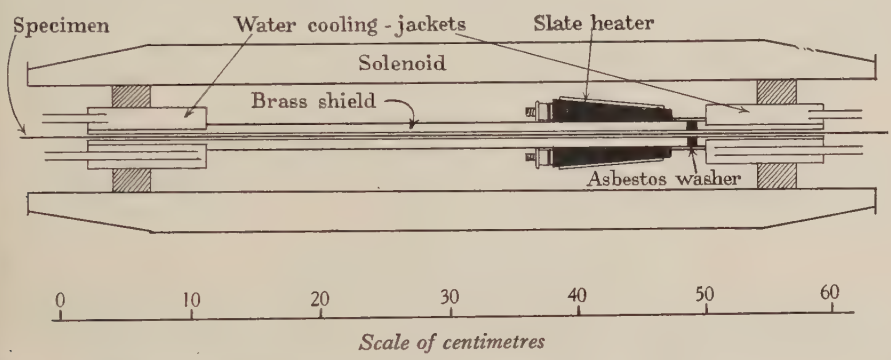


Figure 1. Diagram of the apparatus, showing the temperature-control system in position along the axis of the solenoid.

Temperatures were recorded entirely by movable thermocouples sliding in the bore of the tube. The same technique as that described in the previous paper was employed, and the same remarks concerning precision also apply.

The e.m.f. was measured on a specially adapted potentiometer of thermocouple type manufactured by the Cambridge Instrument Co. Ltd. As supplied this instrument gives readings only to 10μV., and is not sensitive enough for our purpose. A shunt and series resistance were added, as in the *K*-type potentiometer, to increase sensitivity. The resistances were non-inductively wound double-silk-covered Eureka wire no. 28 (for the shunt) and no. 40* (for the series resistance); the resistance of the shunt was made exactly one-nineteenth of the total potentiometer resistance, and that of the series resistance was adjusted so that when both were present in the circuit there was no change in the working current. These adjustments were made by means of a *K*-type potentiometer with very good precision.

* British standard gauge.

A single-pole double-throw switch enabled us to use the potentiometer either in its original form or in its new form; it gave excellently consistent results both ways, and in the sensitive position estimates down to $0{\cdot}1\,\mu$V. were easily and reliably made. The scale reading was $0{\cdot}5\,\mu$V. per division, which is even better than that given by the *K*-type potentiometer. The potentiometer was standardized against a Weston standard cell in a thermostat. A Cambridge high-sensitivity suspended-coil galvanometer was needed to indicate the potentiometer balance, the galvanometer ordinarily used not being sensitive enough to match the increased sensitivity of the potentiometer scale.

With this high sensitivity great care had to be taken with all binding posts, switches, etc., and all electrical connections were separately shielded in grounded tubes. The ice available was not pure enough to provide perfectly constant temperatures, and the ends of the specimen were therefore taken to a bath of good insulating oil where they were connected to the potentiometer leads. The oil was well stirred, and the bath was surrounded by a good heat-insulating case. The junctions were about 2 cm. apart in the oil; no corrosive action was found between the oil and the metals.

§ 3. RESULTS

Figure 2 shows the temperature maps for five typical cases. Table 1 shows the analysis of these five curves, where

$$A = \Sigma\,(dT)^2/dx = \Sigma G.dT, \qquad A$$

$$B = \Sigma\,(dT)^3/dx^2 = \Sigma G^2.dT. \qquad B$$

The values of a and b are found to be

$$a = -1{\cdot}80.10^{-9}\ \text{V.-cm.}/(^\circ\text{C.})^2,$$

$$b = \ \ 1{\cdot}85.10^{-11}\ \text{V.-cm.}^2/(^\circ\text{C.})^3,$$

and the values of $(aA+bB)$ in the table are the theoretical values of the e.m.f. calculated from these values of a and b. The last column shows the measured values of the e.m.f. Positive e.m.f. is directed from the steep gradient to the small gradient in the specimen, which is the direction of the e.m.f. in copper.

Table 1

Curve	$A.10^{-2}$	$B.10^{4}$	$aA+bB$ (μV.)	E (μV.)
1	24·20	7·75	−2·9	−3·0
2	39·37	18·95	−3·6	−3·5
3	41·71	19·94	−3·8	−3·9
4	44·55	19·50	−4·4	−4·5
5	69·93	43·95	−5·2	−5·2

The effect produced by changes in tension has been observed but not examined in detail.

The effect produced by magnetization is considerable. Figure 3 shows e.m.f. plotted against magnetizing field for the five temperature-distributions of figure 2.

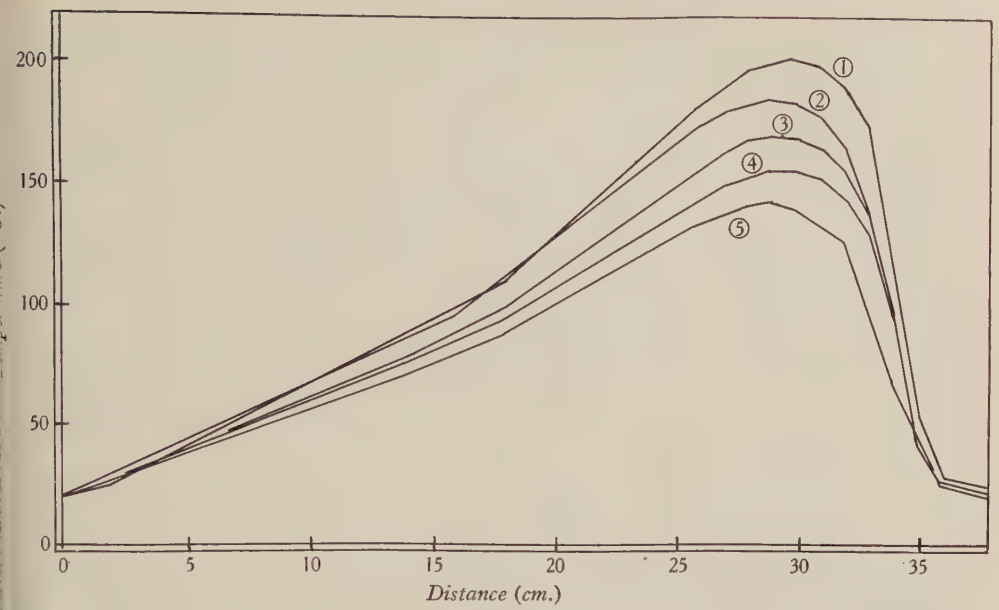


Figure 2. Typical temperature maps, the distributions of temperature are numbered to correspond with the values of e.m.f. recorded in table 1 and figure 3. The curves are drawn as split up for analysis into rectilinear sections.

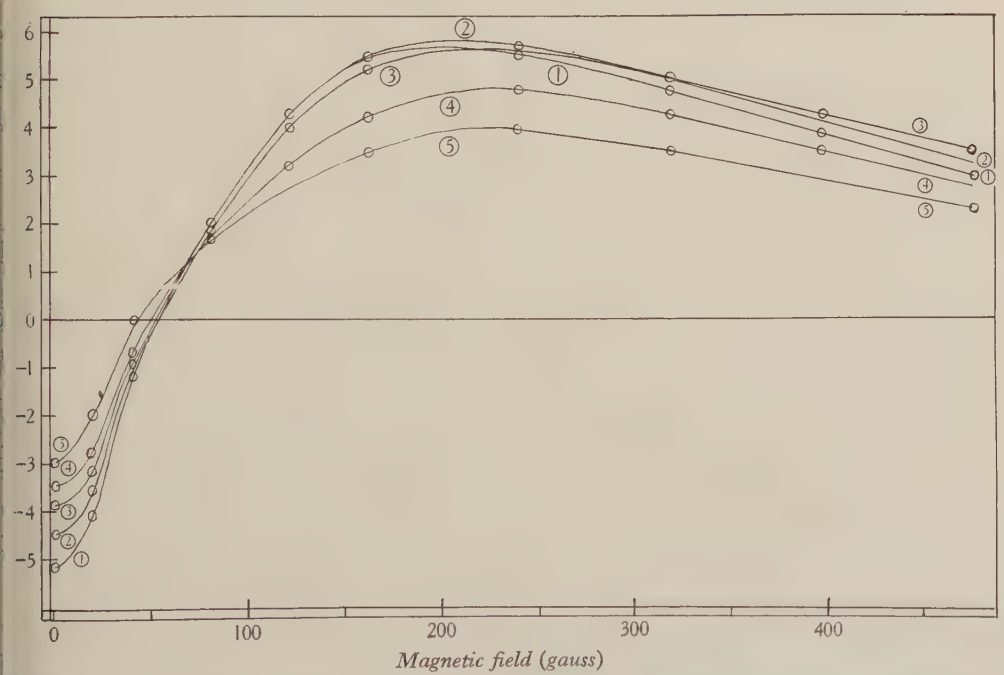


Figure 3. E.m.f. against magnetic field. The circles represent observed values of e.m.f. for ten field-strengths. The curves are numbered to denote corresponding temperature-distributions in figure 2.

The initial e.m.f. in zero field is always negative, but a small magnetic field quickly reverses the sign. At fields around 200 gauss the e.m.f. reaches a maximum positive value of around 5 μV. and thereafter decreases linearly towards zero. We wish to find how the constants a and b vary with magnetic field; each value of the magnetic field was therefore associated with all the temperature-distributions and the corresponding values of e.m.f., and the constants were computed for each field. The agreement between the values of $(aA+bB)$ and the experimental values of e.m.f. was just as good as for the case of zero field shown in table 1. The results for the values of a and b as functions of magnetizing field H are shown graphically in figure 4.

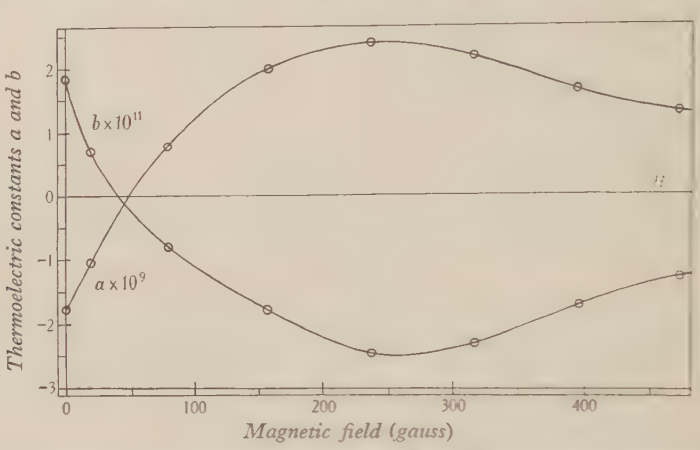


Figure 4. Thermoelectric constants against magnetic field. The circles represent the results of calculations for all temperature maps obtained at the various values of magnetizing field.

Each constant undergoes a reversal of sign at almost 40 gauss, rises to a maximum value of opposite sign at 200 gauss, and thereafter decreases asymptotically toward zero.

§ 4. DISCUSSION

The success of the foregoing analysis makes it evident that there can have been no recrystallization of the specimen to introduce spurious non-homogeneity at uncontrollable points in it. A test was also made for recrystallization by displacing the wire longitudinally through the heating-system; no permanent change of e.m.f. was caused, and this fact indicated real homogeneity of the wire throughout the temperature-ranges used, which were under 200° C.

As in the case of copper the values of a and b have been assumed to be independent of temperature because they have been taken as constant throughout the whole temperature-distribution. In the present work another assumption not immediately obvious has been made: the magnetization of the specimen under the uniform field is not perfectly uniform, owing to changes of susceptibility with temperature, but we have assumed the values of a and b to be constant throughout the specimen even under magnetization. Since both a and b are found to vary

with magnetization this assumption cannot be exactly true. On the other hand the success of the analysis indicates the assumption to be at least approximately true. Further, the greatest temperature-differences existing in the wire are less than 200° C., and nowhere is the temperature near the Curie point for nickel; the susceptibility varies most at the maximum temperature, where the temperature curves are nearly symmetrical and temperature-variations contribute least to the e.m.f. This however is a difficulty which any work of higher precision than the present would have to overcome.

The curves of figure 3 should be compared with the curves of thermomagnetic e.m.f. obtained for nickel by Broili[2], and by Tao and Band[3]. In Broili's work the magnetic field extended over only one temperature-gradient, temperature was kept constant through the varying magnetic field, and the return gradient was well out of the magnetic field. Broili stated that his e.m.f. was independent of the gradient but was a function only of the temperature-difference existing in the uniform magnetic field. This evidently means that Broili's e.m.f. was caused entirely by magnetic modification of the Thomson coefficient t. This Thomson e.m.f. would mask any smaller effect due to changes in a and b in Broili's apparatus. In our present work, since the whole temperature-distribution is within the field, the Thomson coefficient, even though changed by the magnetic field, contributes nothing since the end temperatures are equal.

This last statement is again subject to the following qualification: owing to temperature differences the magnetization of the specimen is not uniform, and hence the value of t is variable in the specimen even for a uniform magnetizing field. In our case the Thomson e.m.f. will still integrate to zero, but it is fruitful to discuss whether differences in magnetization of this kind are responsible for thermomagnetic phenomena.

Thus in the work of Chang and Band[4] it was suggested that the thermomagnetic e.m.f. was directed from parts of the specimen (in that case iron) with higher spontaneous magnetization to parts with lower spontaneous magnetization. The e.m.f. in nickel has the opposite direction. Now if there is an e.m.f. due merely to differences in longitudinal magnetization, we must remember that there are the same differences of magnetization running in reverse directions as the wire emerges from the magnetic field at either end (in the one case at a low temperature, and in the other at a high temperature, if we consider Broili's arrangement). These differences are much greater than the small differences due to the dependence of susceptibility on temperature within the magnetic field; in fact the effect of temperature is merely to destroy the symmetry of the magnetization-distribution. In the arrangement of the present work there is also this asymmetry of magnetization-distribution, but we can hardly consider the idea analytically.

Evidently thermomagnetic e.m.f. is more simply explained by magnetic modification of the three thermoelectric constants; the influence of differences of magnetization due to changes of susceptibility with temperature is probably confined to the smaller hysteresis phenomena discussed in the work of Chang and Band already referred to above.

Finally we point out that our magnetic fields are not strong enough to eliminate the effect of ferromagnetization as distinct from spontaneous or true magnetization probably the changes in *a* and *b* shown in figure 4 are due entirely to ferromagnetic changes or reversals of the elementary magnetic axes of crystals. Whether there will remain some effect on *a* and *b* after saturation of the reversal process cannot be determined without fields of the order of 3000 gauss, as used by Englert in his study of resistance-changes in nickel[5].

REFERENCES

(1) Feng, P. C. and Band, W. The longitudinal thermoelectric effect: (I) copper. *Proc. phys. Soc.* **46**, 515–22 (1934).
(2) Broili, Heinz. Ferromagnetismus und elektrische Eigenschaften: (V). *Ann. Phys.*, Lpz., **14**, 259–72 (1932).
(3) Tao, S. C. and Band, W. Some thermomagnetic effects in nickel and iron. *Proc. phys. Soc.* **44**, 166–8 (1932).
(4) Chang, W. Y. and Band, W. Thermomagnetic hysteresis in steel. *Proc. phys. Soc.* **45**, 602–9 (1933).
(5) Englert, Erno. Ferromagnetismus und elektrische Eigenschaften: (VI). *Ann. Phys.*, Lpz., **14**, 589–612 (1932).

537.322.2:546.621

THE LONGITUDINAL THERMOELECTRIC EFFECT: (3) ALUMINIUM

By M. K. LI, M.S. (Yenching) and WILLIAM BAND, M.Sc.,
Professor of Physics, Yenching University, Peiping, China

Received October 6, 1934. *Read in title May* 3, 1935

ABSTRACT. The thermoelectric e.m.f. produced in aluminium wires is measured by a Paschen galvanometer. The temperature-distribution is analyzed by means of the formula previously used, but without success; alternative formulae also fail. The e.m.fs. observed were less than $0{\cdot}3\,\mu$V.; they were in the same direction as those previously found in copper wire.

§ 1. INTRODUCTION

The e.m.f. in commercially pure aluminium wire produced by asymmetrical heating to temperatures always under 180° C. has been observed and measured. The temperature-controlling system is the same as that used in the work of Feng and Band[1] on copper, but the e.m.f. was so minute that a Paschen galvanometer (made by the Cambridge Instrument Co. Ltd.) had to be employed to detect it. The e.m.f. was in the same direction as that for copper, directed from the steep gradient to the small gradient within the specimen; it was never more than $0{\cdot}3\,\mu$V.

The Paschen galvanometer was calibrated by means of the modified thermocouple potentiometer as described in the paper by Pi and Band[2], which gave results that could be read to $0{\cdot}5\,\mu$V., and estimated to $0{\cdot}1\,\mu$V.* The deflection sensitivity of the galvanometer was adjusted and maintained at $0{\cdot}0032\,\mu$V./mm. at the scale-distance of 1 m.

Even with this high sensitivity, the e.m.f. values obtained are satisfactorily steady during the mapping of the temperature-distribution by means of the thermocouple probe[1]; the mean deviation of the e.m.f. value during probing was usually between 1 and 2 per cent.

The aluminium specimen was connected at its ends to two thick copper leads, the connexions being made in a temperature bath containing good machine oil; it is more important that the junctions should be at identical temperature than that they should be at some definitely known temperature. Leakage and corrosion, and impurities, in an ice bath caused quite uncontrollable irregularities in the e.m.f.

The thick copper leads passed through grounded shields and were connected directly to the galvanometer. A switch would make the circuit impracticably complicated, owing to contact charges. The resistance of the specimen being small

* The potentiometer measured the standard voltage across a standard resistance (made by the Hartmann Braun Co.) and a tapping on this was applied to the galvanometer.

compared with that of the galvanometer, the latter is well short-circuited; it was found that several hours were needed for the galvanometer reading to return to its open-circuit value after the circuit had been completed. Throughout the work this circuit was left undisturbed; the zero reading was determined before the investigation of each temperature map by leaving the whole system at room-temperature overnight, and was found to remain excellently constant for the whole period of several months during which it was used.

§ 2. RESULTS

Eight or ten different temperature-distributions were examined for each of the four different tensions due to loads of 0, 1, 2 and 3 kg. on the ends of the wire which had a diameter of 1·219 mm.

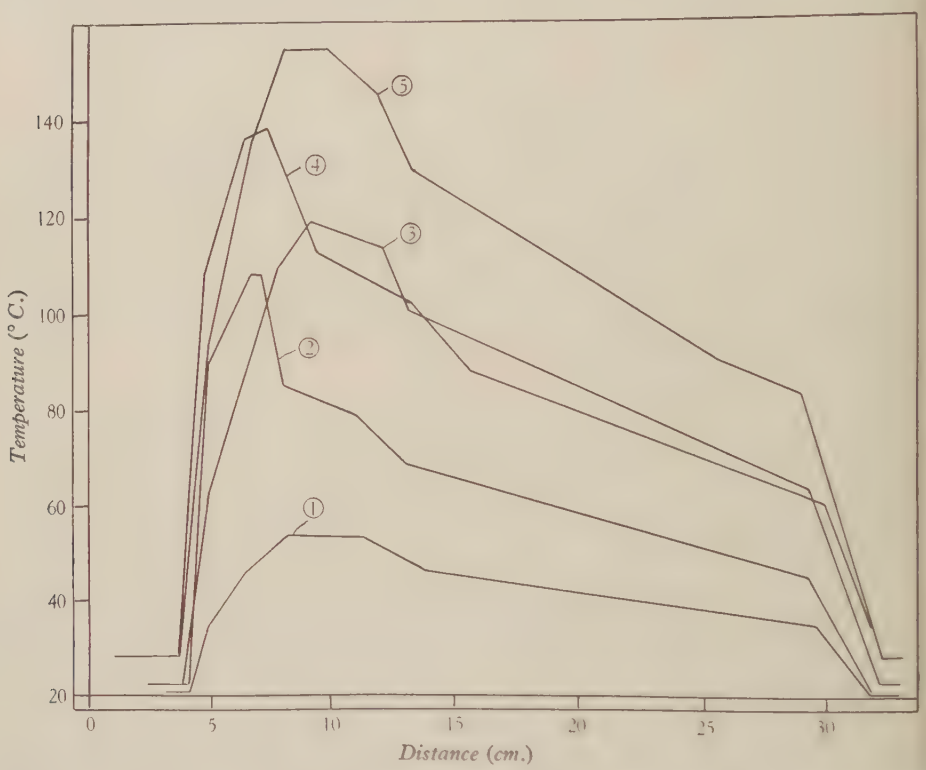


Figure 1. Temperature maps: wire under zero tension. The thermo-e.m.fs. (μV.) were as follow (1) 0·0320; (2) 0·0454; (3) 0·2426; (4) 0·0662; (5) 0·2525.

Typical temperature maps are shown in the figure. These were analysed by th same method as that used for copper, but absolutely no correlation could be foun between the observed e.m.f. values and the integrals from the temperature map It was then thought that if the constants a and b in the formula[1]

$$E = a\int G.dT + b\int G^2.dT$$

were allowed to vary with temperature along the wire, a better correlation might be achieved.

Formulae of the form

$$E = a' \int T.G.dT + b' \int T.G^2.dT$$

and also

$$E = a' \int T.G.dT + b' \int T^2.G^2.dT$$

were used, but with little better success. Not more than three of the maps under any one tension-state could be correlated by any one choice of constants a' and b'.

We conclude that the asymmetrical temperature-distribution is not the controlling factor in the production of the longitudinal thermoelectric e.m.f. in the case of our specimen of aluminium.

There are several possible explanations for the above results. First, what was an apparently uniform and homogeneous wire may have had some impurity non-uniformly distributed in its material; in this case a sudden change in constitution at one point within the high-temperature part would give rise to an e.m.f. independent of the temperature except at that point. Secondly, as a result of previous heating to a high temperature in one part of the wire, the aluminium may have become recrystallized, a discontinuity of structure being thus introduced at one or more points along the specimen. From previous work[3,4,5] it is known that aluminium has recrystallized after being heated for several hours at 500° C., and has done so rapidly at 600° C. In our present work no part of the wire was ever heated above 180° C., and even this temperature was not maintained for more than a few hours at a stretch. It does not seem possible that recrystallization could have taken place, unless at some time in its previous history the wire had been heated to a high temperature in disconnected parts—a highly improbable supposition.

The only fruitful suggestion we can make is that there is an allotropic form of aluminium with a transition temperature somewhere below 180° C., accompanying a structural change so slight that the thermoelectric powers of the two forms differ by not more than about 0·005 μV. per degree of difference in temperature. One of us (W. B.) has succeeded in obtaining a correlation of the present results on the basis of this assumption, but a direct experimental test of the hypothesis will be made before the results of this analysis are published.

REFERENCES

(1) Feng, P. C. and Band, W. The longitudinal thermoelectric effect: (1) copper. *Proc. phys. Soc.* **46**, 515 (1934).

(2) Pi, T. H. and Band, W. The longitudinal thermoelectric effect: (2) nickel, p. 852 of this volume.

(3) Burgers. Recrystallization of aluminium. *Z. Phys.* **59**, 9–10, 651–5 (1930); **67**, 605–78 (1931).

(4) Schmid and Wissermann. Recrystallization of aluminium. *Metallwirtschaft*, **10**, 409–10 (1931).

(5) Karnop and Sarks. Kinetics of recrystallization in copper and aluminium. *Z. Phys.* **60**, 464–80 (1930).

537.322.2:546.621

THE LONGITUDINAL THERMOELECTRIC EFFECT: (4) A FURTHER STUDY OF ALUMINIUM

By WILLIAM BAND, M.Sc., Professor of Physics,
Yenching University, Peiping, China

Received October 12, 1934. *Read in title May* 3, 1935

ABSTRACT. Results obtained in work by Li and Band[1] are analysed by means of the formula[2]

$$E = a\int G.dT + b\int G^2.dT + e.DT,$$

where e is the excess in thermoelectric power of normal aluminium over that of a postulated allotropic form having a transition temperature T_c below 180° C. with delayed reverse transition; DT is the temperature difference between the junctions of the two forms of metal. Excellent agreement for all tensions is found. Under zero tension the values found are

$$a = +0{\cdot}6 \times 10^{-11} \text{ V.-cm./(°C.)}^2,$$

$$b = +2{\cdot}2 \times 10^{-13} \text{ V.-cm.}^2/(\text{°C.})^3,$$

$$e = +0{\cdot}0067\ \mu\text{V./°C.}$$

The same critical temperature, 79 C., is required for all tensions. Direct experimental tests with cooling-curves, a Paschen galvanometer being used to measure the thermo-e.m.f., verify this temperature exactly. All possible explanations of this transition are carefully discussed, the recrystallization theory appearing the most probable. Further work is suggested, and the significance of the present results for methods of increasing precision is pointed out.

§ 1. INTRODUCTION

In foregoing papers under the present general title analyses have been reported of measurements made on copper and nickel by Research Fellows of this Department. The analyses were successfully carried out in both cases on the assumption that copper and nickel were homogeneous metals at all temperatures under 200° C. In the case of aluminium, however, as reported in part III, the analysis was not successful, and the suggestion was there made that a transition temperature existed for aluminium below 180 C.

Here we shall report analyses of the results announced previously in this series, showing that this new suggestion permits of a very satisfactory explanation of all values of the e.m.f. obtained in the experimental work on aluminium.

§ 2. THEORETICAL BASIS

In the preliminary tests made by Li and Band to determine the most suitable sensitivity of the Paschen galvanometer the aluminium wire was heated to about the average temperature subsequently employed. It was found that several hours were necessary before a good steady state was attained. During this time differen

parts of the wire must have been at temperatures ranging between 25° and about 150° C., but the precise distribution of the temperature was not mapped.

The hypothesis which has been taken as the basis of the present analysis is simply that all those parts of the specimen which were, in this preliminary test, heated above a certain critical temperature T_c were changed in form and acquired in consequence a slightly different thermoelectric power from that of the original cold form of the metal. This change of form is supposed to be essentially a low-temperature recrystallization; it will therefore persist for a finite time after the temperature has decreased. Thus there are throughout the rest of the work two discontinuities in the wire at the points P and Q which were at the temperature T_c previously. T_c

Whereas a homogeneous wire will exhibit an e.m.f. E given by the formula E

$$E=a.A+b.B,$$ a, b

where A and B are the integrated values $\int G.dT$ and $\int G^2.dT$, the present hypothesis will modify this formula thus A, B

$$E=a.A+b.B+e.DT,$$

where DT is the temperature-difference between the two points P and Q, and e is the difference between the thermoelectric powers of the two forms of metal. DT

Neither the critical temperature T_c nor the positions of P and Q are known. But in analysis, we note that the maximum temperatures attained in the 1-kilogram tests were greater than those in either the preliminary tests or the zero-load tests. This means that the positions on the 1-kg. curves which are at T_c are the discontinuities for the 2-kg. tests. This gives a limiting factor in the choice of P and Q; in analysing the 2-kg. tests the choice of T_c fixes the positions of P and Q by reference to the 1-kg. tests.

The best method is by trial and error; possible values of T_c are chosen and from these the positions of P and Q and the values of DT are deduced.

If it is then possible to find values of a, b and e to give agreement between observed and calculated values of E, then we have a possible solution of the problem. Considering the variety of the temperature curves and values of e.m.f. observed, the discovery of even a single satisfactory solution would be a very convincing circumstance.

§ 3. RESULTS OF ANALYSIS

The integrals A and B were first worked out as described in the previous papers; they were tabulated with the observed values of E for all loads. Trial values of T_c were then selected, and the quantity DT was found for each temperature map occurring in the 2-kg. tests. The temperatures 130°, 120°, 110°, 100°, 90°, 80° and 75° were used; the positions of P and Q on the 1-kg. curve of highest temperature were then found, and from these the values of DT for each curve of the 2-kg. tests. These were tabulated along with the values of A, B and E.

Simple inspection was enough to eliminate most of these choices; only the

last three gave approximate correlation, and of these, on more careful analysis, only 80° gave more than a few cases of good agreement.

For the next approximation, slight variations were made in the value of T_c, and hence in the positions of P and Q and in the values of DT. Owing to the steepness of the temperature-gradient at one of the points it is impossible to fix that point very precisely by the value of T_c; this permits an independent variation of DT by as much as a few degrees in most cases. It was thought that if within these latitudes a good correlation could be obtained, the choice of constants T_c, a, b, and e would be very effectively limited and therefore rather precisely determined.

Table 1 shows the results which were first worked out on the 2-kg. tests. The values of the constants are given in the third row of table 2, and the value of T_c was found to be 79° for best agreement.

Table 1. 2-kg. load

Curve	$A.10^{-3}$	$B.10^{-4}$	DT	E experimental (μV.)	E theoretical (μV.)
1	0·7	1·7	7	0·048	0·047
2	1·4	5·9	3	0·049	0·047
3	0·9	2·3	15	0·094	0·091
4	1·2	6·9	6	0·050	0·053
5	2·3	7·5	22	0·152	0·155
6	2·6	13·3	10	0·109	0·116
7	4·6	30·5	−6	0·103	0·104
8	4·3	21·8	31	0·265	0·263
9	4·3	28·3	27	0·263	0·263
10	5·9	48·0	0	0·181	0·203

In table 1 and those which follow, positive values of DT refer to cases where the point Q on the small temperature-gradient is at a higher temperature than the point P on the steep gradient.

With these values of T_c, a, b, and e of the 1-kg. and 3-kg. results, analysis was next attempted, and it was found that only reasonable changes in the latter three constants were required to give equally good agreement for the same value of T_c. The analysis of the results for zero tension have not quite as much significance because the positions of P and Q are not fixed by T_c, no previous temperature-distribution having been mapped. P and Q were assumed to be the same as those used for the 1-kg. tests, and good enough agreement was then obtained. The values of the three constants are given in table 2, and the agreement is shown in the following tables.

Table 2

Load (kg.)	$a.10^{11}$	$b.10^{13}$	e (μV.)
0	+0·6	+2·2	+0·0067
1	0·5	2·0	0·0051
2	1·0	3·0	0·0050
3	1·2	3·4	0·0052

In every case, a change of one significant figure in the values given for the three constants destroys the good agreement, unless unreasonably large latitude or error is permitted in the positions of P and Q as determined by the critical temperature T_c.

Table 3. Zero load

Curve	$A.10^{-3}$	$B.10^{-4}$	DT	E experimental (μV.)	E theoretical (μV.)
1	0·3	0·65	4·5	0·032	0·033
2	3·3	20·50	−3·0	0·045	0·045
3	1·9	8·00	30·0	0·243	0·230
4	6·0	54·50	−13·0	0·066	0·068
5	5·1	41·00	23·0	0·253	0·254
6	3·1	18·60	16·0	0·156	0·155
7	5·0	38·00	8·0	0·165	0·167
8	3·8	23·00	13·0	0·159	0·161

Table 4. 1-kg. load

Curve	$A.10^{-3}$	$B.10^{-4}$	DT	E experimental (μV.)	E theoretical (μV.)
1	2·2	10·5	19·0	0·126	0·129
2	2·2	11·0	20·0	0·133	0·135
3	5·7	50·0	−6·0	0·097	0·099
4	1·7	7·8	17·0	0·119	0·112
5	3·6	23·0	4·0	0·085	0·085
6	3·5	24·0	−3·5	0·044	0·044
7	5·2	43·0	23·0	0·229	0·229

Table 5. 3-kg. load

Curve	$A.10^{-3}$	$B.10^{-4}$	DT	E experimental (μV.)	E theoretical (μV.)
1	1·3	5·0	4·0	0·054	0·054
2	2·2	12·0	3·5	0·085	0·084
3	3·2	21·0	−2·0	0·098	0·098
4	1·7	7·3	21·5	0·158	0·158
5	2·5	10·4	42·0	0·288	0·284
6	2·9	20·7	24·0	0·231	0·230
7	5·6	47·0	0·0	0·222	0·226

The load was not increased beyond 3 kg. because at 4 kg. the wire would approach closely to its elastic limit and the specimen became spoilt after a short time.

Positive values of e here obtained indicate that electrons flow in the recrystallized portion (between P and Q) from the hot end to the cold end—i.e. from Q to P, where DT is positive in the above tables. This indicates that the thermoelectric power of the recrystallized metal is greater than that of the normal metal by the value of e.

§ 4. EXPERIMENTAL TEST OF DISCONTINUITY AT 79° C.

The above indirect evidence of a discontinuity dependent upon temperature, which looks like a recrystallization of aluminium at 79°, permits of a simple experimental test. This was carried out as follows, after the foregoing analysis had been completed.

A new sample of the same aluminium wire was taken and mounted in much the same way as for the previous work, except that two thermocouples were attached to the wire at fixed points P and Q and a single travelling heater surrounded the wire between these points. The two copper-constantan couples were of the paper-insulated probe type previously used; they were adjacent to the specimen, and their junctions were bound to it with no. 40 s.w.g. copper wire tied around several times to improve thermal contact. Their free ends were connected with a galvanometer through a key so constructed that they were never connected electrically, except through the specimen.

The same Paschen galvanometer was used to detect the e.m.f. produced and, as before, full shielding was used to protect it against stray fields coming from the resistances controlling the heating-current. Sensitivity was set at about 100 mm. μV., at which value the zero is perfectly stable.

First the whole length of wire between P and Q was heated to about 70° C. for half an hour, and a test was made to see whether nearly symmetrical heating over P alone would produce an e.m.f. Temperature-differences of over 50° C. were tested between P and Q, and any discontinuity at these points of the kind postulated would have caused a measurable e.m.f. No indication of such an e.m.f. was found in this case except small deflections explicable as due to slight disturbances of the symmetry of temperature-distributions.

The process was repeated, P to Q being above 80° C. while P and Q were kept as nearly as possible at 80° C. This condition was maintained for about 15 minutes, and after a few minutes of cooling tests were again made at P and Q with the heater carrying only a small current. When P was at 60° and Q at 30° a definite voltage of the order of 0·2 μV. was found. It vanished when the heater was placed symmetrically between P and Q, and which reversed sign when Q was heated and P allowed to cool.

Next, the heater was placed so that the thermocouple at Q just entered the end of the axis of the heater, and the couple P was shifted and remounted in a symmetrical position at the other end of the axis, say at P'. In this way the whole of $P'Q$ could be heated without moving the heater, while the original point P was not heated by more than a few degrees.

The current was turned on, and what must have been an almost symmetrical temperature-distribution caused a smoothly increasing e.m.f. as the temperatures at Q and P' rose. Both temperatures remained equal to within 0·5° C. The e.m.f. steadily increased until the temperatures reached 78°, when its value remained constant while the temperature rose to 84°. It then suddenly became irregular,

increasing by a deflection of several millimetres and decreasing suddenly to the value which it had at 84°. These jerks continued until the temperature reached 90° when a steady increase again set in. At 95° the heating current was cut off and a curve of cooling was obtained; it was quite smooth except for a small lag at the commencement, and did not show any discontinuity at 80°; see *A* in figure 1.

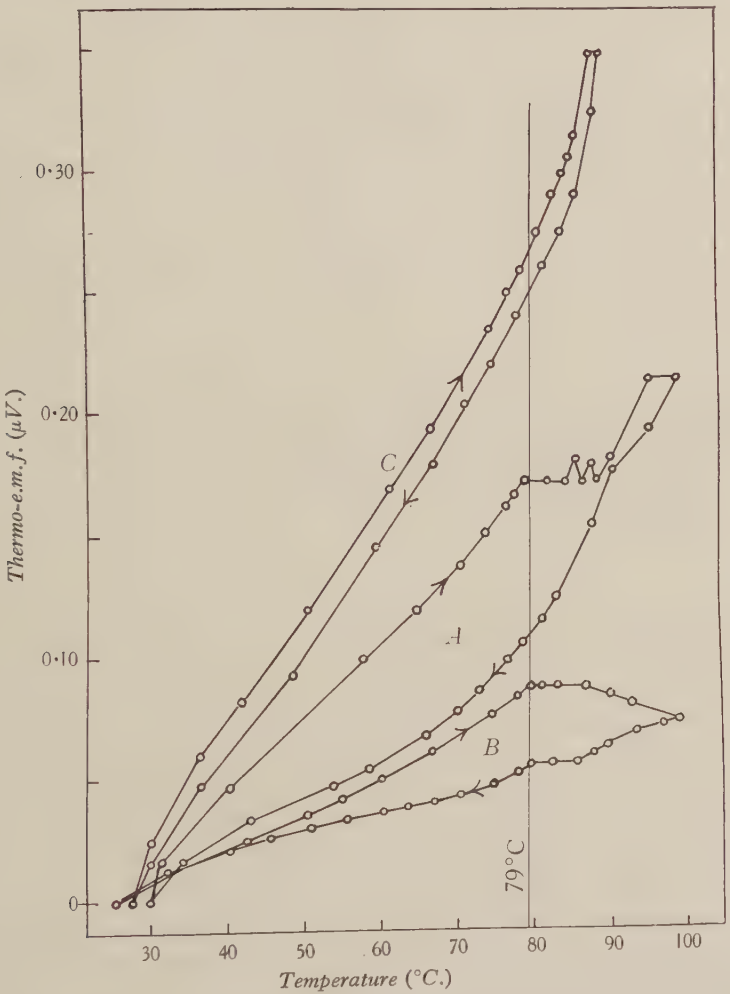


Figure 1.* Hysteresis cycles of thermo-e.m.f., showing the transition temperature at 79° C. (*a*) after 15 minutes local heating at 80° C.; (*b*) after 3 hours local heating at 80° C.; (*c*) after 6 hours entire heating at 80° C.

The part between P' and Q was now heated above 80° symmetrically for 3 hours, the two thermocouples being kept at about 80° throughout that time. After several hours of cooling the heater was shifted so that the point Q came to the middle of

* 12 hours later a cycle similar to *A* was obtained, showing partial recovery from recrystallization by the entire specimen.

the axial bore in the heater, and the cycle shown at B in figure 1 was obtained. Here, as before, the heater carried a current sufficient to give a final temperature around 120°, and the rate of change of temperature with time was of convenient speed and quite steady. Temperature at the point Q and the e.m.f. were recorded every half-minute except at the last part of the cooling curve below 60°. It is to be noted that there was no increase in e.m.f. at temperatures above 79° where the steady value was reached; instead, a decrease was observed which continued when the heating current was cut off at 95°. That this was not due to a shift of the galvanometer zero is shown by the satisfactory return to the same zero after the apparatus had acquired room-temperature again. The whole cycle was completed in 1 hour.

Finally the specimen was completely removed from the apparatus, immersed in an oil bath and kept at a temperature slightly above 80° C. for 6 hours. Retesting at the same points P' and Q as before gave the curve C in figure 1; from the facts that there is no sudden break in the gradient at 79° C., and that the hysteresis loop has almost vanished, it is evident that the uniformity of the wire had been restored.

The fact that after treatment the wire gave a bigger e.m.f. than before is difficult to explain, since the heater was placed symmetrically between the thermocouples just as for the test A. The only directional physical property that could cause any e.m.f. under such symmetrical heating, other than a sudden discontinuity, would be an internal tortional strain. The results of Terado, Tsutsui and Tamano should be compared here[3].

§ 5. DISCUSSION

The change of form at 79° C. The direct tests described above are convincing evidence for a distinct change of form at 79° C., and the simplest explanation of this would be a recrystallization theory. Alternative explanations present themselves. A reaction between the aluminium and some metallic impurity in the wire may cause a heterogeneity at some critical temperature, or else an oxide coating may develop at a fairly definite temperature and produce its own thermo-e.m.f. In the latter case, since the oxide coat would become thicker with time, it is difficult to explain the quite definite value of e found in the analysis; also the temperature at which the oxide formed must depend on the rate at which the metal is brought to that temperature—a perfectly critical temperature is practically impossible.* The other alternative hypothesis would also seem too improbable, because chemical changes would be far too slow at such temperatures, and the reverse change would take place as easily on cooling; again, the actual temperature of transition would depend on the rate of change of temperatures and no true critical value would be possible.

The recrystallization theory seems therefore the most probable; but since commercial aluminium usually contains as much as 0·1 per cent of impurity,† it is

* It is also known that a very thin oxide film exists always at room-temperatures.

† Chiefly graphite and silica.

just possible that recrystallization of a metallic impurity like aluminium carbide is responsible for the present results. It would obviously be desirable to repeat the direct test on 100-per cent pure metal, if such could be obtained in a practically useful form; but Peltier effects between junctions of the pure specimen with the less pure wire connexions would offer a considerable difficulty. Also, to avoid the criticism regarding oxidation, the specimen would have to be *in vacuo*, a provision which would make it almost impossible to avoid local heatings that would destroy the homogeneity of the specimen.

Methods of increasing precision. Assuming that our hypothesis is accepted as a working basis, probably the indirect analysis of the e.m.f. produced by mapped temperature-distributions is the most precise method of obtaining the change of thermoelectric power at recrystallization. It is almost impossible to get a perfectly symmetrical temperature-distribution so as to eliminate the e.m.f. due to the homogeneous effect and thus to isolate the exceedingly small e.m.f. due to contact between the two forms of the metal. Thus the direct test described in this paper, while precise enough for the critical temperature, is not capable of giving more than a rough estimate of e.

In order to get precise determinations of a and b, either all temperatures should be kept below 80°, which would make voltage-measurements more difficult, or else the whole specimen should be annealed at a temperature a little above 80° C. for several hours before each test. Preferably a and b should be found for both forms of the metal separately by both of these methods, and then e could be determined by using temperature maps in which both forms are present.

Throughout the quantitative work reported in this series of papers, reliance has been placed on the movable thermocouple and fixed junction in gentle contact with the specimen. The precision of the probe method has been discussed in part (I); but the present results for aluminium lend further support for that method.

Thus, it might be thought that a direct measurement of the temperature actually within the wire at several points would be a desideratum of precision; this has in fact been suggested to the writer by one competent critic. However, it is at once obvious that any attempt to use solder or welding to get real thermal contact between the junction and the specimen would heat the wire above the critical temperature here discovered, and so ruin the homogeneity of the specimen in an unmeasurable way. Further, any mechanical deformation of the wire required to insert a junction through its surface would again vitiate the claim that the wire is a homogeneous circuit, and would put our work back into the crude constriction class to which most previous work belongs[4].

It is therefore felt even more certainly that the only method of increasing precision is to increase the sensitivity of the voltage-measuring apparatus so that quite small temperatures and temperature-gradients can be used. By properly designed shields the space around the wire could be brought reasonably close to the temperature of the wire itself. While the probe obviously introduces a disturbance, no better method seems available; and such disturbance can be averaged out fairly completely in practice.

§ 6. CONCLUSION

In conclusion one may say that results for aluminium have been analysed and show excellently precise agreement with the homogeneous theory, if we also admit of a discontinuity of structure with a transition temperature at 79° C. The evidence that this discontinuity is due to the element aluminium is not perfectly conclusive until further tests have been made on 100 per cent pure metal. That the discontinuity is in the nature of a recrystallization seems the most probable explanation. It is obvious that ordinary methods of investigation such as recalescence during cooling would fail to detect such a slight change as appears to take place according to the present work. It is also doubtful if X-ray analysis could do so, because that method has given, until recently, somewhat inconsistent results when applied to the much greater recrystallization that occurs above 500° C.[5].

ADDITIONAL NOTE

by Dr W. BAND, received October 10, 1934

The three figures show typical results obtained on similar wire specimens. Symmetrical heating between two points A and B above 80° C. was maintained for several hours, no e.m.f. being observed. Immediately after cooling the heater was shifted over A, and a heating-cooling cycle of e.m.f. against T_A was obtained, figure 2. Immediately after this cycle had been completed, a second heating-cooling

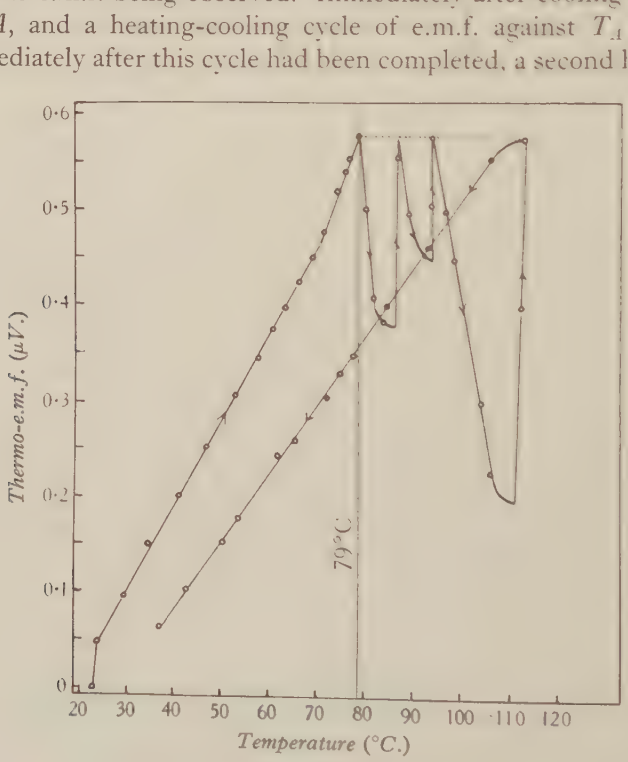


Figure 2.

process gave an entirely different curve above 79°, figure 3. Sample curves (e.m.f. against T_B) obtained when the heater was over B are shown in figure 4.

From these data we can see that the e.m.f. produced in the newly recrystallized parts of the wire depends for its direction upon the direction of the temperature-

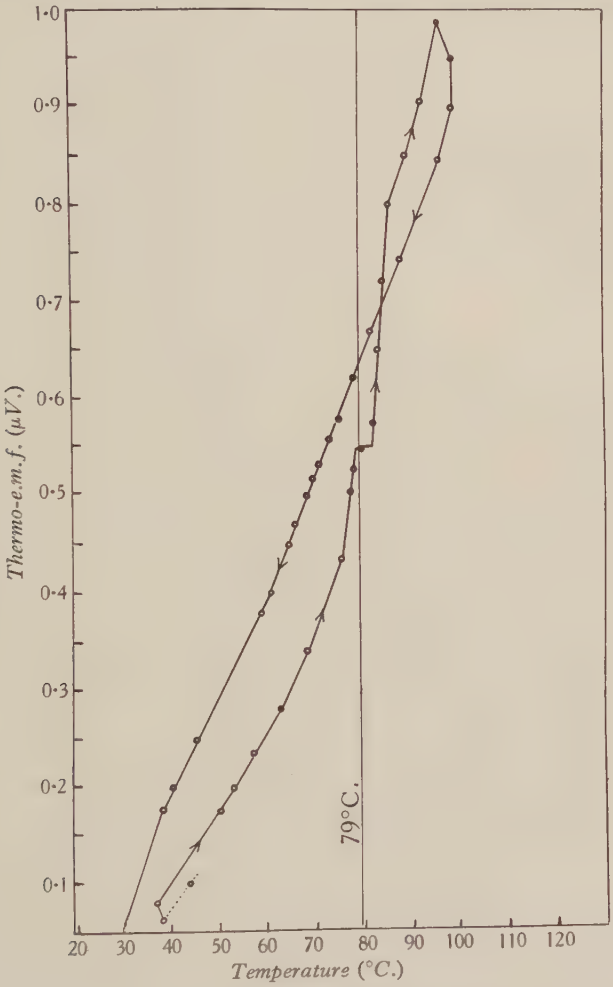


Figure 3.

gradient in those parts, viz.: the difference between figure 2 and figure 4. But it is difficult to explain how a curve like that of figure 3 could follow the cycle shown in figure 2.

The magnitudes of the voltages obtained in these tests seem to depend upon the mechanical treatment of the wire during mounting in the apparatus, and thus presumably upon small internal strains. Such strains would only even out after a considerable time. This would explain why, during the original measurements previously reported, it was found necessary to maintain a steady temperature-distribution for 4 hours or more before a steady e.m.f. could sometimes be obtained.

That the so-called transition at 79° C. may be connected with slight re-adjustments of internal strain rather than with recrystallization is an interesting possibility. It is felt, however, that further speculation had better await the results of more refined methods of investigation suggested by the experiences of the present exploratory work.

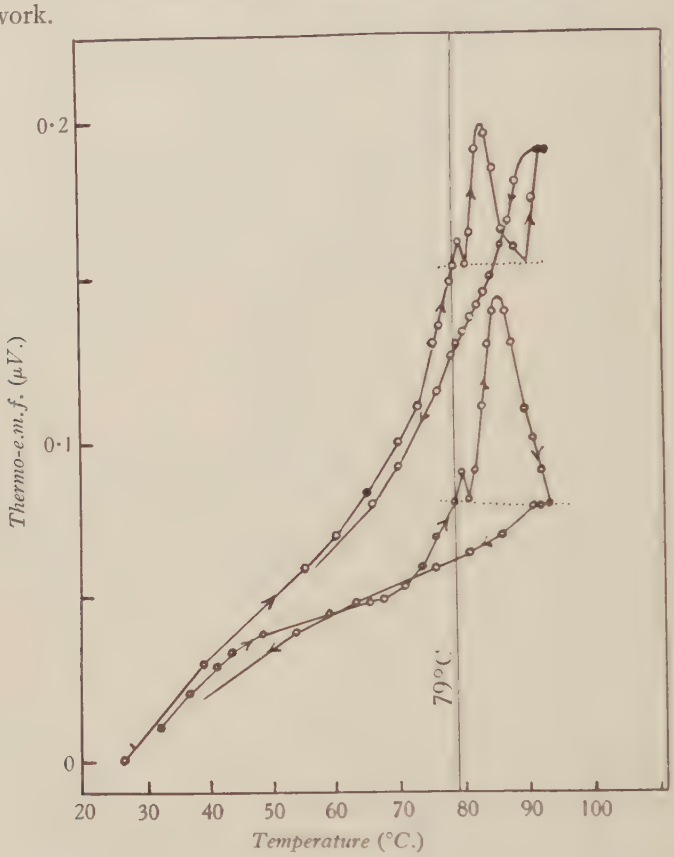


Figure 4.

REFERENCES

(1) Li and Band. The longitudinal thermoelectric effect: (3) aluminium, p. 859 of this volume.

(2) Li and Band. The longitudinal thermoelectric effect: (1) copper. *Proc. phys. Soc.* **46**, 515 (1934).

(3) Terado, Tsutsui and Tamano. *Sci. Pap. Inst. phys. chem. Res.*, Tokyo, **7**, Nos. 122 and 195.

In this connexion one would like to be more certain that a symmetrical temperature-distribution can be moved along a wire and still remain symmetrical; temperature waves are almost sure to commence with the motion of the heater, and these would at once cause the e.m.f. fluctuations observed by these workers. There seems no particular need to postulate fluctuations of internal crystal-arrangement.

(4) W. Hume-Rothery, in his *The Metallic State*, pp. 107–114 (Oxford 1931), gives an excellent critical summary of work done previously to that date.

(5) Schmid and Wassermann. *Metallwirtschaft*, **10**, 409–10 (1931).

539.17

THE DISINTEGRATION OF BORON BY NEUTRONS

By H. J. TAYLOR, M.Sc.
Professor of Physics, Wilson College, Bombay

Received May 11, 1935. *Read June* 7, 1935

ABSTRACT. The tracks of the particles which result from the disintegration of boron by neutrons have been recorded by impregnating the emulsion of a photographic plate with borax. It is shown that there are two modes of disintegration, corresponding to the reactions:

$$B^{10} + n^1 \rightarrow Li^7 + He^4,$$
$$B^{10} + n^1 \rightarrow He^4 + He^4 + H^3.$$

Photomicrographs of the tracks, showing both modes of disintegration, are reproduced.

§ 1. INTRODUCTION

A BRIEF report has already been published[1], recording the discovery of tracks corresponding to the disintegration of boron and lithium when bombarded by neutrons. The purpose of the present paper is to give a fuller statement of the results obtained in the case of boron. The tracks of the fast particles which are liberated in the nuclear reactions are recorded directly in a photographic emulsion. The technique employed has been fully described in a previous paper[2].

§ 2. METHOD OF EXPERIMENT

Briefly stated, the method consists in impregnating the emulsion of a photographic plate with a suitable salt of the element to be examined. The plate is then exposed to the neutrons, and developed in the usual way. The tracks of the heavy particles which result from the disintegration are registered in the emulsion; each track is visible, under a high magnification, as a row of developed grains in a straight line. For satisfactory work special emulsions free from background and of very fine grain are required. Ilford R plates have been specially prepared to meet these requirements, and have been used for the present experiments.

The source of neutrons for most of the experiments consisted of a quantity of radon sealed in a tube with powdered beryllium. The initial amount of radon varied for different experiments, the maximum being about 200 millicuries. Some experiments were also done with a polonium-beryllium source, but this was found less satisfactory on account of the comparative weakness of the polonium sources available. Exposure to fast neutrons was effected by placing the plate as near as possible to the source, but with some 6 cm. of lead interposed, in order to reduce the direct effect of the γ rays on the plate. To expose the plate to slow neutrons, both

source and plate were surrounded by a large quantity of paraffin. Under these conditions, however, fast neutrons are still present in the radiation reaching the plate. Conversely, some slow neutrons are present even in the absence of paraffin, on account of scattering which takes place in the table and other surrounding objects.

§ 3. THE DISINTEGRATION OF BORON

Boron is introduced into the emulsion by soaking the plate in a strong solution of borax and allowing it to dry in the dark. It is only possible to make a rough estimate of the amount of boron thus introduced, but it is of the order of 0·05 mg. per cm^2 of plate-surface. This is only 1 per cent of the amount required, for slow neutrons, to reduce the intensity of the radiation to one half of its initial value[3]. The treatment of the plate with borax was found not to affect the photographic properties of the emulsion seriously. The borax was generally washed out before development, but this seems unimportant.

Disintegration by slow neutrons. With slow neutrons the probability of disintegration is relatively high, and exposures of the order 1000 millicurie-hours were found to be ample. The best plates show as many as 50,000 tracks per cm^2, sensibly straight, and with the number of grains per track varying up to ten. The mean length of the tracks is $7{\cdot}6\mu$. Photomicrographs of typical tracks are shown in figures 1 and 2. Most of the tracks are very clear and well defined, but it is difficult to determine the equivalent range in air exactly, for the following reasons. (i) The measured length (i.e., the distance between the extreme grains of a track) shows, as one would expect, large fluctuations owing to the random distribution of grains in the emulsion. (ii) The true range will in general exceed the measured length by an unknown amount, since the end grains need not be at the extreme points of the range. (iii) The conversion factor to reduce the range in gelatine to the range in air is subject to some uncertainty. The uncertainties can, however, be partly allowed for on the basis of previous experience[2], and the best value for the equivalent range in air which can at present be deduced from the measurements is

$$1{\cdot}1 \pm 0{\cdot}1 \text{ cm.}$$

These tracks must be interpreted as being due to the disintegration of the boron nucleus into two particles. As the momentum of the neutron is small these particles move in sensibly opposite directions, giving rise to an apparently single track.

With our present knowledge of nuclear masses, the only reaction which appears to fit these observations is

$$B^{10} + n^1 \rightarrow Li^7 + He^4 \qquad \text{.....(1).}$$

The release of energy, corresponding to the observed range, is 2×10^6 electron-volts.

Until recently, the accepted masses of the neutral atoms of Li^7 and He^4 and of the neutron have been as follows: $Li^7 = 7{\cdot}0146$, $He^4 = 4{\cdot}0022$, $n^1 = 1{\cdot}0080$. With these values reaction (1) requires the value $B^{10} = 10{\cdot}011$, disagreeing markedly with Aston's value $B^{10} = 10{\cdot}0135$.

Recently, however, the available evidence has been reviewed[4, 5, 6], and it

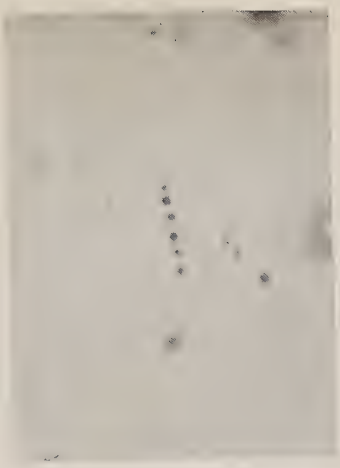


Figure 1.

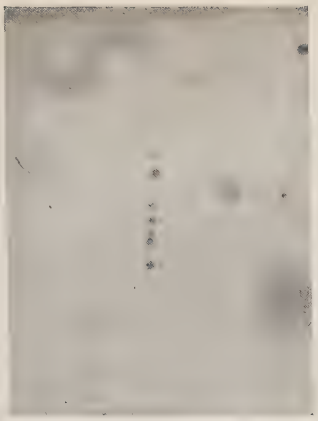


Figure 2.

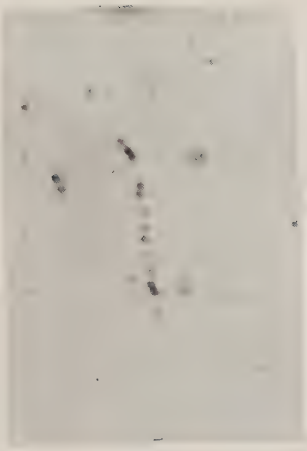


Figure 3.

Figure 4.

Figure 5.

(1) and (2) Typical examples of tracks resulting from reaction (1), and due to a slow neutron.
(3) Tracks corresponding to reaction (1), but due to a fast neutron.
(4) Three tracks corresponding to reaction (2). The long H^3 track rises out of the plane of focus.
(5) The same photographed in a higher plane, showing the H^3 track to better advantage.
Magnification of all the photographs: $\times 1930$.

appears that the accepted masses may require some modification. At the moment he most probable values are[4]

$$n^1 = 1{\cdot}0083$$
$$He^4 = 4{\cdot}0034$$
$$Li^7 = 7{\cdot}0170$$
$$B^{10} = 10{\cdot}0143.$$

These masses fit reaction (1) within the limits of error.

Disintegration by fast neutrons. Boron under fast neutron bombardment gives many disintegrations of type (1), some of which are due to the slow neutrons always present in the radiation. In some cases the tracks of Li^7 and He^4 can be separately seen, as the angle between them is less than 180°. This is an indication that a fast neutron, with considerable momentum, has been responsible for the disintegration. Figure 3 shows an example of this.

Several examples of what appears to be a different type of disintegration have been found. The best example is reproduced in figures 4 and 5. Three tracks diverge from a single point, two shorter ones denoted by *a* and *b*, and a longer one denoted by *c*. Track *c* does not lie in the plane of the photograph, but rises steeply from the point of disintegration. The horizontal component of *c* may be denoted by *c'*.

The ranges and angles are as follows:

$a = 0{\cdot}7$ cm. in air, angle a, $b = 102°$
$b = 0{\cdot}5$ cm. in air, angle b, $c' = 120°$
$c = 3{\cdot}0$ cm. in air, angle c', $a = 138°$
angle c, $c' = 35°$.

The angle c–c' in the vertical plane is subject to an uncertainty of about 5°.

One might advance the hypothesis that two of the tracks, *a* and *c*, represent a disintegration of type (1), whilst the third track is independent and situated at this point purely by chance. On this hypothesis the resultant momentum of the two particles can be worked out, and this must be the same as the momentum of the incident neutron. The energy of the incident neutron calculated in this way is found to be 13×10^6 eV., of which half must have been absorbed in the reaction. It is evident that the hypothesis breaks down, even when allowance is made for very considerable errors in the measurements. There can thus be little doubt that we are really dealing here with a disintegration into three particles. It appears that the only reaction which agrees with the observations is

$$B^{10} + n^1 \rightarrow He^4 + He^4 + H^3 \qquad \ldots\ldots(2).$$

With the masses recently proposed[4] ($H^3 = 3{\cdot}0161$), it will be seen that the reaction requires an energy of 3×10^5 eV. to be supplied, which explains why it does not occur with slow neutrons. In this example the combined energy of the three particles is $3{\cdot}3 \times 10^6$ eV., so that the energy of the incident neutron must have been $3{\cdot}6 \times 10^6$ eV.

The resultant momentum of the three particles can be calculated, but the uncertainty in the angle c–c' affects the value obtained. This resultant momentum

should be the same as that of the incident neutron, and within the limits of accuracy of the measurements no discrepancy is found.

Several other examples of what are apparently similar cases of triple disintegration have been found, but no other which is so well placed for measurement. This type of disintegration is clearly much rarer than type (1), the relative frequencies of occurrence, in the plates exposed to fast neutrons, being of the order 1 : 1000.

§ 4. ACKNOWLEDGMENTS

My thanks are due to Mr Goldhaber, who suggested some months ago that lithium and boron might be amenable to investigation by this method, and who has also helped in making the exposures. I also desire to thank Lord Rutherford and Dr Chadwick of the Cavendish Laboratory, where the experiments have been done, for their interest and encouragement. I am greatly indebted to Mr Bloch and his staff, of the Ilford Company, for producing the special photographic material. Finally, I must express my thanks to the Senatus of Wilson College, Bombay, for extended leave of absence.

REFERENCES

(1) Taylor and Goldhaber. *Nature*, Lond., **135**, 341 (March 2, 1935).
(2) Taylor. *Proc. roy. Soc.* **150**, 382 (1935).
(3) Amaldi, D'Agostino, Fermi, Pontecorvo, Rasetti, and Segrè. *Proc. roy. Soc.* **149**, 522 (April 10, 1935).
(4) Oliphant, Kempton, and Rutherford. *Proc. roy. Soc.* **150**, 241 (May 1, 1935).
(5) Bethe. *Phys. Rev.* **47**, 633 (April 15, 1935).
(6) Aston. *Nature*, Lond., **149**, 522 (April 10, 1935).

535·375·54:538.221

A NOTE ON THE RAMAN SPECTRUM OF A FERROMAGNETIC OXIDE

By L. F. BATES, D.Sc., Ph.D., Reader in Physics,
University College, London

AND

H. E. HOGWOOD, B.Sc.

Received April 4, 1935.

ABSTRACT: An attempt has been made to determine whether the ferromagnetic oxide Cr_5O_9 exhibits a Raman spectrum. No trace of such a spectrum was found.

In recent years our knowledge of the ferromagnetism of metals has been considerably increased, but our knowledge of the nature of ferromagnetism in the metallic oxides is still very meagre. Now, it occurred to one of us (L. F. B.) that valuable information might be obtained if it were possible to examine the Raman spectra of such oxides, for changes in the Raman spectrum with rise in temperature might show whether a particular grouping of atoms was responsible for the observed ferromagnetic properties, and might also indicate the nature of the coupling between the atoms. A perusal of the literature showed that the oxide of chromium, Cr_5O_9, would probably be the best substance with which to make an attempt. The magnetic properties of this oxide have been studied by Dr H. Sachse[1], who found that it possesses a ferromagnetic Curie point at about 108° C. We are deeply indebted to Dr Sachse and to Fräulein Lorentz for the gift of samples of Cr_5O_9, which we were able to compare with those prepared for our work.

The oxide, Cr_5O_9, was prepared by passing a stream of chromyl chloride vapour over hot glass or silica maintained at a temperature of about 370° C. Suitable heavy deposits were usually obtained in a few hours, although thin, translucent deposits were formed on the inner walls of tubes in a few minutes. The preparation of crystals of Cr_5O_9 as large as those obtained by Dr Sachse requires much time and patience, especially as glass becomes brittle after being in contact with hot chromyl chloride vapour for some time. The oxide shows a pronounced violet surface colour and appears green by transmitted light. It is difficult to determine the thickness of a layer unless it is deposited on a previously weighed piece of glass, for it is extremely difficult to remove.

In a preliminary survey the light from a powerful quartz mercury vapour lamp was concentrated upon and passed through a thin oxide deposit formed on the

inner surface of a bulb of transparent silica, about 2 cm. in diameter, so as to enter a quartz Raman spectrograph of great light-collecting power, manufactured by Messrs Hilger. Direct light was carefully excluded, and with an exposure of 150 hours it was found that little beyond the 3132 lines was transmitted, and no new lines appeared.

A more thorough survey, particularly in the ultra-violet, was made by concentrating the light upon a thick coating of the oxide inside a tube which had been cut obliquely at one end, the light falling upon the projecting portion. The light was then scattered from the deposit many times before it reached the collimator slit. The result of a 100-hour exposure to the scattered light was compared with that of a 10-sec. exposure to the direct beam. The chief difference between the two spectrograms was the appearance of some fluorescence on the long-wave sides of the more intense lines, but no evidence of new or Raman lines was found.

This method of attack was less satisfactory for the visible region of the spectrum, where the following procedure was adopted. Heavy deposits of Cr_3O_9 were made on two strips of fused silica 12 cm. long and 2·5 cm. wide. These were mounted parallel, with the deposits facing each other, but so displaced laterally that the incident light could be concentrated upon the one plate. After being scattered several times, the light was finally scattered from the other plate into the spectrograph. The result of 300 hours' exposure with this arrangement, when compared with that obtained with one hour's exposure with the deposits covered with white blotting paper, again showed no Raman lines to be present.

It is therefore concluded that no Raman lines can be obtained in experiments on light scattered from the upper layers of this ferromagnetic oxide. It must be recorded, however, that the experiments here described were made under very adverse conditions of scattering, and the main purpose of this note is to report the attempts made, in the hope that other workers may possibly discover a ferromagnetic compound more suitable for Raman investigations.

In conclusion, we desire to thank Professor E. N. da C. Andrade for the facilities placed at our disposal, especially for the loan of the Raman spectrograph, and Mr H. Terrey for help in preparing the oxide deposits.

REFERENCE

(1) H. Sachse, *Zeit. für Phys.* **70**, 546 (1931).

548·73

THE ABSORPTION FACTOR FOR THE POWDER AND ROTATING-CRYSTAL METHODS OF X-RAY CRYSTAL ANALYSIS

By A. J. BRADLEY, D.Sc., Royal Society Warren Research Fellow

Communicated by Prof. W. L. Bragg, F.R.S., June 17, 1935.

ABSTRACT. A method has been devised for calculating the absorption factor for the powder method. It may be used for all values of μr, where μ is the linear absorption coefficient and r the radius of the specimen. It is, however, especially suitable when μr exceeds 2. For smaller values of μr, for which the calculations are tedious and become less accurate, it is better to use Claassen's graphical method. By such a combination of the two methods a set of data was calculated for which the absorption factor may easily be obtained for any value of μr and for any angle of reflection, with an error not exceeding 1 per cent. The data may also be applied for the calculation of the absorption factor in the lines belonging to the layer of zero order in rotation photographs.

§ 1. INTRODUCTION

In the Debye-Scherrer method of crystal analysis a beam of X rays bathing a cylindrical specimen of powdered crystalline material gives rise to a diffraction pattern. The relative intensities of the lines in this pattern are very much influenced by the absorption of the X rays in the specimen, but hitherto no entirely satisfactory method has been devised for calculating the numerical value of the absorption factor. Solutions for particular cases have been given by different authors[1,2,3], while Claassen[4] has devised a graphical method which is of general application. Rusterholz[5] has shown that for very highly absorbing specimens it is possible to calculate the absorption factor, which is in this instance a fairly simple function of the glancing angle θ.

It is, however, possible to calculate the approximate absorption factor by a new method which gives results accurate to at least 1 per cent, except for very small absorptions where the Claassen graphical method is entirely satisfactory. The chief importance of the new method for calculating the absorption factor is that it gives accurate results in the region where Claassen's method is less reliable. Results are here tabulated for all values of μr, where μ is the linear absorption coefficient of the specimen and r is its radius. Only a limited number of angles have been considered, but the absorption factor for intermediate angles is quickly found by graphical interpolation.

For either a single crystal or a mass of crystalline powder the intensity of reflection, in the absence of absorption, is proportional to the total volume V of the

specimen which is immersed in the incident X-ray beam. Thus for a cylindrical h specimen the intensity is proportional to $\pi r^2 h$, where h is the height of the specimen in the beam. This statement does not hold good where there is finite absorption of X rays in the specimen, the contribution of different portions of the specimen being no longer equivalent. In many cases the absorption effects are so great that only a small portion of the periphery of the specimen makes any appreciable contribution towards the intensity of reflection. The contribution of a tiny fragment of crystal of volume dV is reduced by the fraction $e^{-\mu a}$, where μ is the linear absorption coefficient of the whole specimen and a is the total length of path of the ray through the specimen before and after reflection from the fragment. In order to evaluate the total strength of the reflection, the value of $e^{-\mu a} dV$ must be integrated throughout the volume of the specimen. The resulting expression then replaces the factor V in the intensity formula. For a cylindrical specimen $\pi r^2 h$ is replaced by the expression $h\iint e^{-\mu a} d\sigma$, where $d\sigma$ is a small element of the cross-section of the specimen, and the integration is carried out over the whole cross-section. The intensity of the reflection is, therefore, cut down by absorption in the ratio

$$A = \frac{1}{\pi r^2} \iint e^{-\mu a} d\sigma.$$

A is termed the absorption factor. It may be expressed in the alternative form

$$A = \iint e^{-\mu r x} ds \qquad \text{.....(1),}$$

where $x = a/r$ and $ds = d\sigma/\pi r^2$.

To evaluate equation (1) it is necessary to obtain the value of $ds\,dx$ for all possible values of x. The first step (described in § 2) is to calculate the value of s, the proportion of the cross-section of the specimen for which the absorbing path xr is less than a given value—e.g. the area $XSX'Y$ in figure 1. This gives s as a function of x, which may be differentiated to give $ds\,dx$. The problem is made more complicated because the same formulae are not applicable to all portions of the cross-section of the specimen. In figures 3 to 8 different portions are distinguished according to the formulae which are applicable. The arguments used in § 2 are based primarily on figure 1, which is typical of the front areas of figures 3 to 8. Figure 2 is typical of a back area. No direct method could be found for calculating the corresponding formulae for the centre areas. Suitable corrections can, however, be made for the latter. Except for small absorption, the formulae deduced for the front areas are the most important, and for large values of μr the rest may be disregarded.

The calculations in § 2 give s as an ascending-power series in x, which is convergent for $x \not> 2$. This is quite satisfactory since, as will be seen later, x cannot exceed 2 in the front areas of figures 3 to 8. Claassen has suggested that such a series might be used to calculate the absorption factor for large values of μr in the following way.

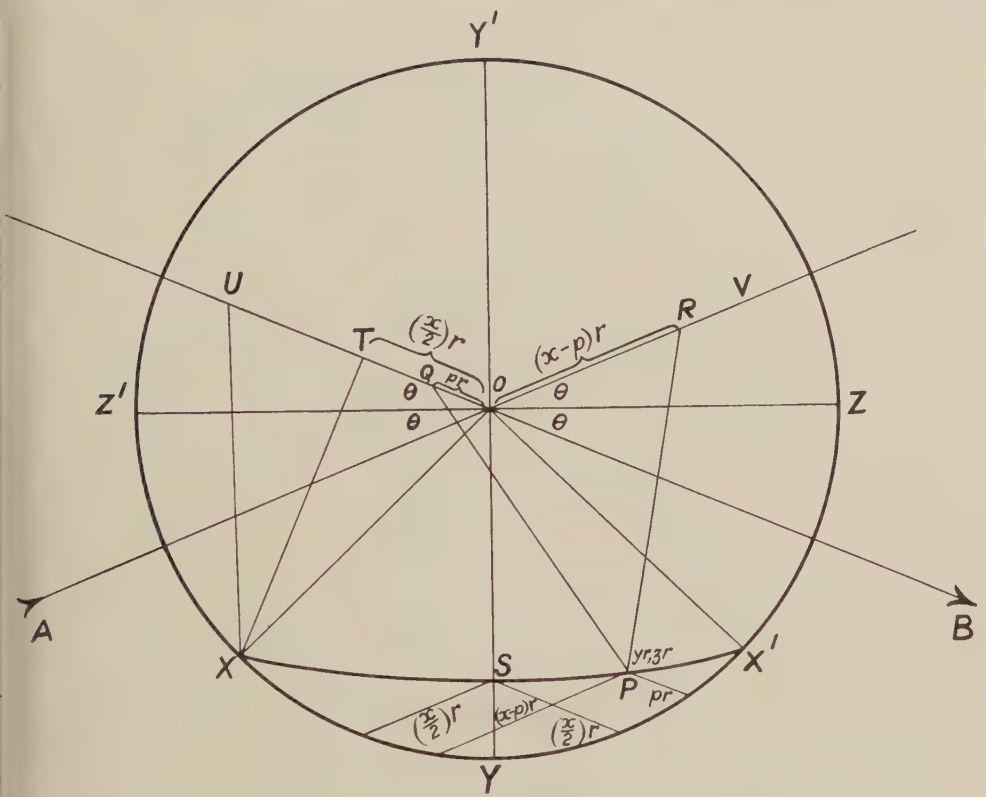
Putting $$s = \alpha x + \beta x^2 + \gamma x^3 + \ldots \qquad \text{.....(2),}$$

we find that $$\frac{ds}{dx} = \alpha + 2\beta x + 3\gamma x^2 + \ldots \qquad \text{.....(3).}$$

Then according to Claassen equation (1) for large values of μr gives approximately

$$A = \int_{x=0}^{x=\infty} \frac{ds}{dx} e^{-\mu r x}\, dx = \frac{\alpha}{\mu r} + \frac{2!\beta}{\mu^2 r^2} + \frac{3!\gamma}{\mu^3 r^3} + \ldots \qquad \ldots\ldots(4).$$

It will be found, however, that if one substitutes for α, β, γ, etc. the coefficients obtained in § 2, the above series becomes divergent for all values of μr and is therefore invalid. This difficulty is avoided by replacing the upper limit $x = \infty$ by the exact maximum value of x for the front portion of the specimen. Then, unfortunately,



General case

Figure 1.

equation (4) becomes considerably more complicated, and since the maximum value of x is to some extent a function of θ it is necessary to consider separately the cases of reflection at different angles.

The value of A found from a series such as that in equation (2) represents the contribution of either the front or the back portion of the specimen. For the centre portions an estimate must suffice, since it is not possible to obtain a series for them. The complete absorption factor is the sum of the values for the front, back and centre portions. However, the contribution from the front portion usually far outweighs the others, which are really in the nature of correction terms.

§ 2. THE AREA OF THE SPECIMEN FOR WHICH THE ABSORBING PATH IS LESS THAN A GIVEN VALUE

In figure 1 AO and OB indicate respectively the directions of the incident and reflected beams. The incident beam is supposed to be parallel and completely to bathe a cylindrical specimen of radius r. θ is the glancing angle. YOY' and ZOZ' are respectively normal and parallel to the reflecting crystal planes. XSX' is the locus of all reflecting points for which the sum of the incident and emergent paths

Y, Z
X

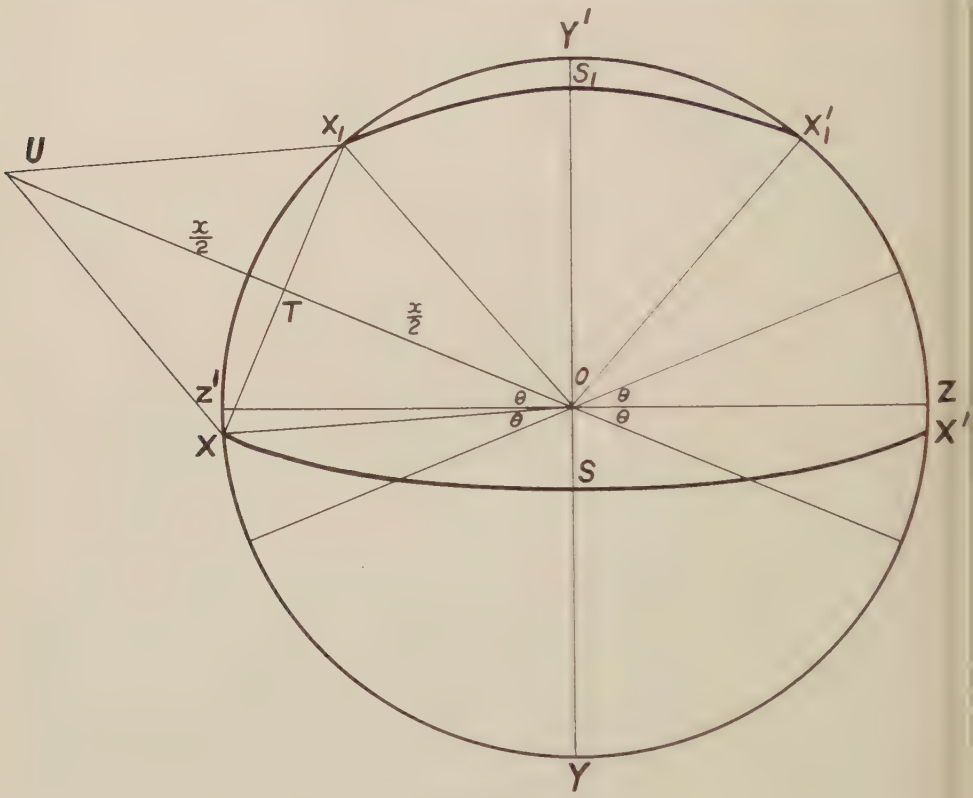


Figure 2.

is equal to xr. For a ray reflected from S the incident and emergent paths are each equal to $xr/2$. Let P be any point along XSX' such that the emergent path is pr and hence the incident path is $(x-p)\,r$.

p

According to Claasen, the position of the point P is given by the following construction. Along AO produced mark off OR equal to $(x-p)\,r$. Along BO produced mark off OQ equal to pr. Then $QP = RP = r$, the radius of the specimen. The position of all other points along XSX' may be found by varying the value of OQ between zero and xr, while that of OR is varied between xr and zero.

It is required to find the area of $XSX'Y$, which is that portion of the specimen in which all particles are situated for which the sum of the incident and reflected

paths is less than x. It is first necessary to determine each of the areas $OXYX'$ and $OXSX'$ separately.

The area $$OXSX' = r^2 \int_{z=0}^{y^2+z^2=r^2} y\,dz - r^2 \int_{z=0}^{y^2+z^2=r^2} z\,dy \qquad \text{.....(5).}$$

It is necessary to evaluate y and z in terms of a common parameter, for which purpose the following relation is used:

$$PQ = PR = r.$$

The y, z coordinates of Q are $-p\sin\theta$, $-p\cos\theta$.

R are $-(x-p)\sin\theta$, $(x-p)\cos\theta$.

Hence
$$\left.\begin{aligned} PQ^2 &= r^2(y+p\sin\theta)^2 + r^2(z+p\cos\theta)^2 = r^2 \\ PR^2 &= r^2\{y+(x-p)\sin\theta\}^2 + r^2\{z-(x-p)\cos\theta\}^2 = r^2 \end{aligned}\right\} \qquad \text{.....(6).}$$

On subtraction, $$x(2p-x) + 2y(2p-x)\sin\theta + 2zx\cos\theta = 0,$$

$$z = (1-2p/x)\tan\theta\,(\tfrac{1}{2}x\operatorname{cosec}\theta + y).$$

Let $$(1-2p/x)\tan\theta = q \text{ or } p = \tfrac{1}{2}x(1-q\cot\theta) \qquad \text{.....(7).}$$

Then $$z = q(\tfrac{1}{2}x\operatorname{cosec}\theta + y) \qquad \text{.....(8).}$$

Substituting for z and p in equation (6), we have

$$\{y+\tfrac{1}{2}x(1-q\cot\theta)\sin\theta\}^2 + \{q(\tfrac{1}{2}x\operatorname{cosec}\theta + y) + \tfrac{1}{2}x(1-q\cot\theta)\cos\theta\}^2 = 1,$$

$$y^2(1+q^2) + xy(1+q^2)\sin\theta + (\tfrac{1}{2}x)^2(1+q^2) - 1 = 0,$$

$$y = -\tfrac{1}{2}x\sin\theta \pm \sqrt{\{1/(1+q^2) - (\tfrac{1}{2}x)^2\cos^2\theta\}}.$$

Let $$\pm\sqrt{\{1/(1+q^2) - (\tfrac{1}{2}x)^2\cos^2\theta\}} = t \qquad \text{.....(9).}$$

Then $$y = -\tfrac{1}{2}x\sin\theta + t \qquad \text{.....(10),}$$

and $$z = q(\tfrac{1}{2}x\cos^2\theta\operatorname{cosec}\theta + t) \qquad \text{.....(11).}$$

For the locus XSX', x is constant, and p (q) is the variable. Hence

$$\frac{dy}{dq} = \frac{dt}{dq}; \quad \frac{dz}{dq} = t + q\frac{dt}{dq} + \frac{x}{2}\cos^2\theta\operatorname{cosec}\theta,$$

$$y\frac{dz}{dq} = -\frac{x}{2}\sin\theta\frac{dz}{dq} + t^2 + qt\frac{dt}{dq} + \frac{x}{2}t\cos^2\theta\operatorname{cosec}\theta,$$

$$z\frac{dy}{dq} = qt\frac{dt}{dq} + \frac{x}{2}q\cos^2\theta\operatorname{cosec}\theta\frac{dt}{dq},$$

$$\int y\,dz = -\tfrac{1}{2}x\sin\theta\int dz + \int t^2dq + \int qt\,dt + \tfrac{1}{2}x\cos^2\theta\operatorname{cosec}\theta\int t\,dq,$$

$$\int z\,dy = \int qt\,dt + \tfrac{1}{2}x\cos^2\theta\operatorname{cosec}\theta\int q\,dt.$$

The required area may now be found from equation (5), inserting the appropriate limits. For the case shown in figure 1, we have

at S, $p = \tfrac{1}{2}x$, $q = 0$; at X', $p = 0$, $q = \tan\theta$.

Hence Area $$OXSX' = r^2\int_{q=0}^{q=\tan\theta}(y\,dz - z\,dy)$$

$$= r^2\left[-\tfrac{1}{2}x\sin\theta\int_{q=0}^{q=\tan\theta}dz + \int_{q=0}^{q=\tan\theta}t^2dq + \tfrac{1}{2}x\cos^2\theta\operatorname{cosec}\theta\int_{q=0}^{q=\tan\theta}(t\,dq - q\,dt)\right] \qquad \text{.....(12).}$$

The above integral can be expanded from equations (9), (10), (11). The ambiguity in equation (9) leads to two solutions. Since figure 1 shows a positive value for y, the positive sign must be taken for t. This rule applies throughout the front areas of figures 3 to 9. For the back areas of figures 2, 3 and 4, y is always negative. These areas correspond to negative t in equation (9). The two cases will be considered separately, and finally a reference will be made to the centre areas of figures 3 to 8 which do not fall into either class.

Case 1. *Front Areas only, as in figure* 1. From equations (9), (10), (11), (12), taking the positive sign in equation (9),

$$\text{Area } OXSX' = r^2\left[-(\tfrac{1}{2}x)^2 \sin 2\theta - \tfrac{1}{2}x\sqrt{\{1-(\tfrac{1}{2}x)^2\}} + \theta + x\cos^2\theta \operatorname{cosec}\theta \int_{q=0}^{q=\tan\theta} \sqrt{\{1\ (1+q^2) - (\tfrac{1}{2}x)^2\cos^2\theta\}}\, dq\right] \quad \ldots\ldots(13a).$$

Area $OXYX'$ is found from figure 1 as follows:

Let $$OU = xr;\ OT = \tfrac{1}{2}xr.$$

Then X, the end of the locus XSX', is given by

$$UX = OX = r.$$

Since T is the midpoint of OU, XTO is a right-angled triangle. Hence

$$\cos T\hat{O}X = \tfrac{1}{2}x,$$

$$X\hat{O}S = \pi/2 - X\hat{O}Z' = \pi/2 - X\hat{O}T + Z'\hat{O}T = \sin^{-1}(\tfrac{1}{2}x) + \theta.$$

Hence area $$OXYX' = r^2 X\hat{O}S = r^2\{\sin^{-1}(\tfrac{1}{2}x) + \theta\} \quad \ldots\ldots(14a).$$

From equations (13*a*) and (14*a*), in figure 1

$$\text{Area } XSX'Y = \text{area } OXYX' - \text{area } OXSX'$$

$$= r^2\left[\sin^{-1}(\tfrac{1}{2}x) + (\tfrac{1}{2}x)^2 \sin 2\theta + \tfrac{1}{2}x\sqrt{\{1-(\tfrac{1}{2}x)^2\}} - x\cos^2\theta \operatorname{cosec}\theta \times \int_{q=0}^{q=\tan\theta} \sqrt{\{1/(1+q^2) - (\tfrac{1}{2}x)^2\cos^2\theta\}}\, dq\right] \quad \ldots\ldots(15a).$$

Case 2. *Back Areas only, as in figure* 2. From equations (9), (10), (11), (12), taking the negative sign in equation (9),

$$\text{Area } OX_1S_1X_1' = r^2\left[-(\tfrac{1}{2}x)^2 \sin 2\theta + \tfrac{1}{2}x\sqrt{\{1-(\tfrac{1}{2}x)^2\}} + \theta - x\cos^2\theta \operatorname{cosec}\theta \int_{q=0}^{q=\tan\theta} \sqrt{\{1\ (1+q^2) - (\tfrac{1}{2}x)^2\cos^2\theta\}}\, dq\right] \quad \ldots\ldots(13b).$$

This formula gives a negative value for the area, because it lies completely on the negative side of the x axis.

Area $OX_1Y'X_1'$ from figure 2 by the same construction as that used in figure 1 gives

$$X_1\hat{O}S_1 = \pi/2 - X_1\hat{O}Z' = \pi/2 - X_1\hat{O}T - Z'\hat{O}T = \sin^{-1}(\tfrac{1}{2}x) - \theta.$$

Hence area $$OX_1Y'X_1' = r^2 X_1\hat{O}S_1 = r^2\{\sin^{-1}(\tfrac{1}{2}x) - \theta\} \quad \ldots\ldots(14b).$$

From equations (13*b*) and (14*b*) we may deduce the area $X_1S_1X_1'Y'$ in figure 2.

Since the area given by equation (13*b*) is a negative quantity, it is necessary to change the signs before subtraction.

$$\text{Area } X_1S_1X_1'Y' = \text{area } OX_1Y'X_1' - \text{area } OX_1S_1X_1'$$

$$= r^2\left[\sin^{-1}(\tfrac{1}{2}x) - (\tfrac{1}{2}x)^2 \sin 2\theta + \tfrac{1}{2}x\sqrt{\{1-(\tfrac{1}{2}x)^2\}} - x\cos^2\theta \operatorname{cosec}\theta \times \int_{q=0}^{q=\tan\theta} \sqrt{\{1/(1+q^2) - (\tfrac{1}{2}x)^2\cos^2\theta\}}\, dq\right] \quad \ldots\ldots(15b).$$

It will be seen that equations (15*a*) and (15*b*) differ only in the sign of the second term. These equations may be expanded in the form of a power series in x. With the exception of the term $\pm(\frac{1}{2}x)^2 \sin 2\theta$ the series contains only odd powers of x.

The expansion for s is found by dividing by πr^2. In order that the unique character of the second term may be indicated in equations (15*a*) and (15*b*) this term is placed first in the following series:

$$s = \pm\frac{x^2}{\pi}[\tfrac{1}{4}\sin 2\theta] + \frac{x}{\pi}\left[1 - \frac{\log(\sec\theta+\tan\theta)}{\sec\theta\tan\theta}\right] + \frac{x^3}{\pi}\left[-\frac{1}{3}\cdot\frac{1}{2^3} + \frac{1}{2^3}\cos^2\theta\left\{\frac{1}{2} + \frac{1}{2}\frac{\log(\sec\theta+\tan\theta)}{\sec\theta\tan\theta}\right\}\right] + \frac{x^5}{\pi}\left[-\frac{1}{5}\cdot\frac{1}{2^2}\cdot\frac{1}{2^5} + \frac{1}{2^2}\cdot\frac{1}{2^5}\cos^2\theta\left(\frac{1}{4} + \frac{3}{4}\cos^2\theta\left\{\frac{1}{2} + \frac{1}{2}\frac{\log(\sec\theta+\tan\theta)}{\sec\theta\tan\theta}\right\}\right)\right] + \ldots\ldots \text{ to infinity} \quad \ldots\ldots(16a) \text{ and } (16b).$$

The positive sign for the first term in equation (16) applies to the front of the specimen (16*a*) and the negative sign to the back of the specimen (16*b*). The above series are convergent unless x exceeds 2. It will be shown in the next section that x never exceeds 2 in the front and back portions of the specimen.

Case 3. *Centre Areas.* If figure 1 be modified to correspond to reflection from a centre area, it is found that the locus XSX' does not always terminate on the circumference of the circle. In such cases the limits of integration in equations (5) and (12) are clearly invalid. In other cases equation (5) is valid but not equation (12). In all these cases series (16) is invalid. An alternative series cannot be found except for the case of $\theta = 90°$, to which special consideration will be given. In other cases, the centre areas are relatively unimportant. Indeed for all values of μr exceeding 5 no correction need be applied for them.

§ 3. CHANGE OF ANGLE

It is necessary to define the areas within which equation (16) is valid at different angles. In figures 3 to 8 the areas indicated by the words "front" and "back" both conform to this equation, but the sign of the term $\pi^{-1}(\frac{1}{2}x)^2 \sin 2\theta$ is positive for the front area and negative for the back area. The essential conditions for both front and back areas are that the locus XSX' in figure 1 should terminate on the

circumference of the circle at points X and X', at which points $q = \pm \tan\theta$. The limits of integration which lead to equation (16) depend on this assumption. In figure 3, which corresponds to $\theta = 0°$, all the loci are parallel to AB and the back is merely a replica of the front. The maximum path is $2r$, so that x does not exceed 2 and there is no central area.

In figures 4 and 5, corresponding to $\theta = 22\frac{1}{2}°$ and $\theta = 30°$ respectively, the boundary of the central area is marked by the loci corresponding to $x = 2$. Inside this area the loci fail to reach the circumference and the conditions of figure 1 are not

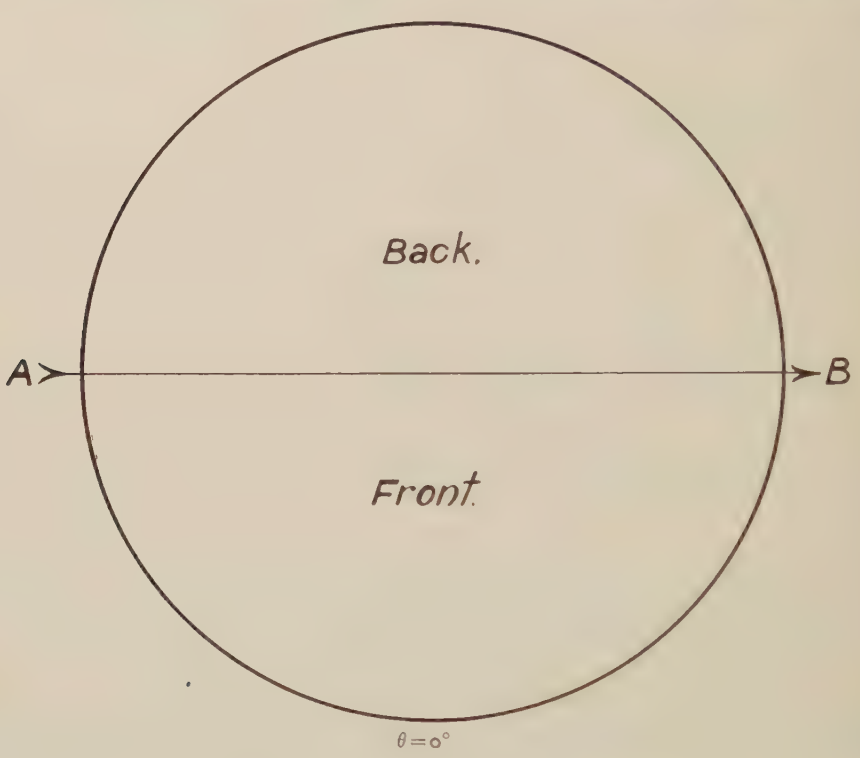


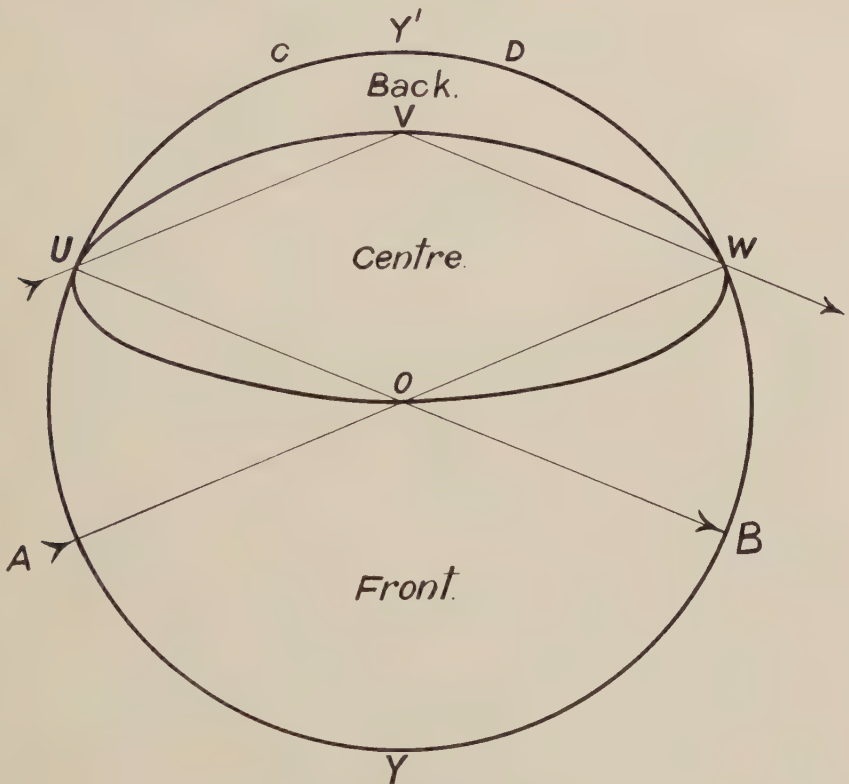
Figure 3.

fulfilled. In figure 5 a small area is shaded. Here the loci reach the circumference but are not terminated at $q = \tan\theta$, so that again equation (16) is inapplicable, although here x is somewhat less than 2.

In figure 6 ($\theta = 45°$) the boundary of the central area is marked by the locus UOW, corresponding to $x = 2$. The incident ray is tangential at U and the reflected ray is tangential at W. Thus from $\theta = 0°$ to $\theta = 45°$ the boundary of the front area is given by $x = 2$, but beyond 45° this is no longer true. The boundary locus is now displaced towards lower values of x, as may be seen from figure 7 ($\theta = 67\frac{1}{2}°$). The boundary locus is now UTW, which is fixed by the condition that the incident ray is tangential at U and the reflected ray tangential at W. It is clear that x is less than 2 for the locus UTW, but at all higher values of x the incident ray must intersect

the specimen before reaching its point of reflection, so that the condition $p=0$, $q=\tan\theta$ is invalidated. At 90°, figure 8, the whole of the specimen is included in the central area, since it is impossible to satisfy the condition $q=\tan\theta$ except on the locus $x=0$.

From $\theta=45°$ to $\theta=90°$ the boundary locus between the front and centre areas is $x=2\sin 2\theta$. To understand what happens if x exceeds this value we may refer



$\theta=22\frac{1}{2}°$

Figure 4.

to figure 6 ($\theta=45°$). A locus such as MNP ends at M and P on the circumference. At M and P the following equation is satisfied:

$$y^2+z^2=1,$$

where y and z have the meaning of equation (6). From the above equation together with equation (6) it can be shown that

$$\{xy\sin\theta+(\tfrac{1}{2}x)^2\}\{q^2\cot^2\theta-1\}=0.$$

Thus either (*a*) $q=\pm\tan\theta$, the limit of integration in equation (12);

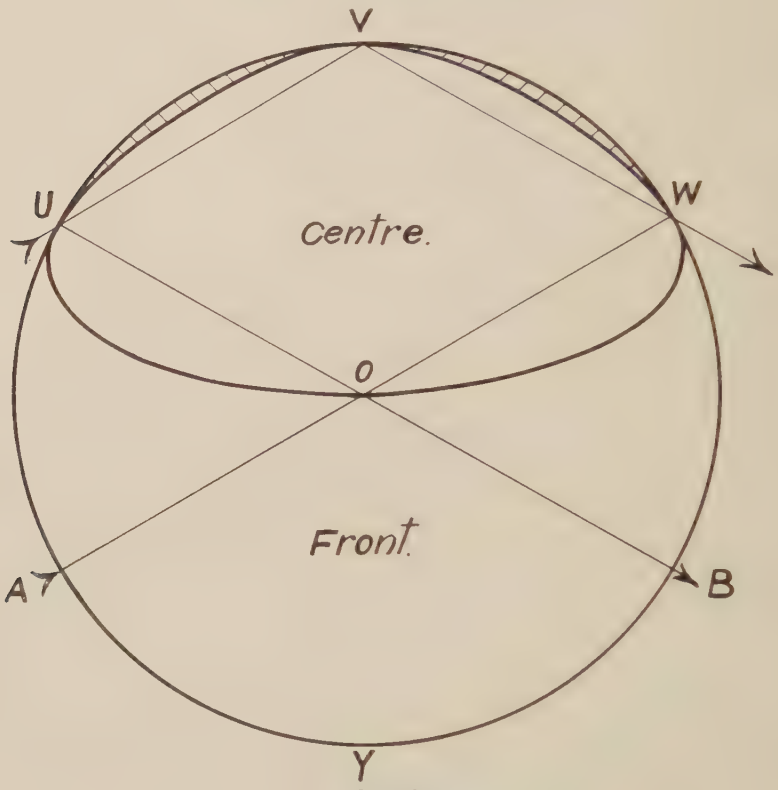
or (*b*) $y=-\frac{1}{4}x\operatorname{cosec}\theta$, whence $q=\pm\sqrt{\{(4/x)^2\sin^2\theta-1\}}$.

Solutions (*a*) and (*b*) are equivalent if $x=2\sin 2\theta$, i.e. at U and W, in either figure 6 or figure 7. It is therefore clear that condition (*a*), which holds at the end of a locus

in the front region, is replaced by condition (*b*) on crossing from the front region to the centre region.

It is not possible to deduce an alternative to equation (16) which can be applied to the central area except for $\theta = 90°$. Equation (16) for 90° would give

$$r^2 [\sin^{-1} \tfrac{1}{2}x + \tfrac{1}{2}x\sqrt{\{1-(\tfrac{1}{2}x)^2\}}].$$



$\theta = 30°$

Figure 5.

Replacing the limit $q = \tan\theta$ at X in figure 1 by the limit appropriate for the central area, $q = \pm\sqrt{\{(4-x)^2 \sin^2\theta - 1\}}$, we obtain the following value for the area *MNPY* in figure 8:

$$\text{Area } MNPY = r^2 [2 \sin^{-1} \tfrac{1}{4}x + \tfrac{1}{2}x\sqrt{\{1-(\tfrac{1}{4}x)^2\}}] \qquad \ldots\ldots(17).$$

This formula replaces equation (16) for the special case in which $\theta = 90°$, and may, of course, be expanded in a similar manner.

Table 1. Limiting values of x for front, back and centre areas

θ	Front		Back		Centre	
	Min.	Max.	Min.	Max.	Min.	Max.
0°–30°	0	2	4 sin θ	2	2	2 sec θ
30°–45°	0	2	—	—	2	2 sec θ
45°–90°	0	2 sin 2θ	—	—	2 sin 2θ	4 sin θ

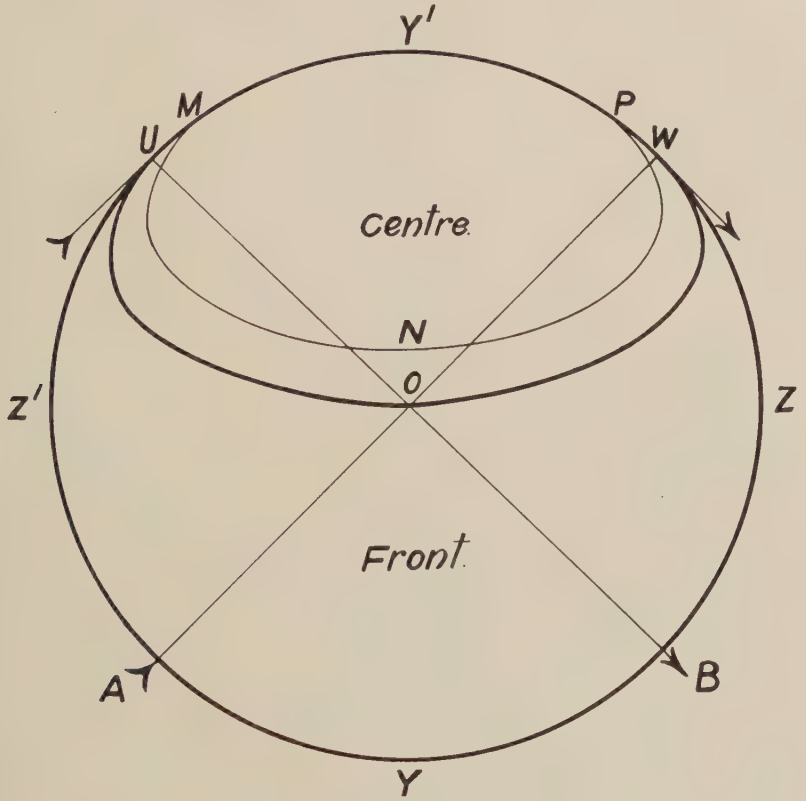
§ 4. THE ABSORPTION FACTOR: CONTRIBUTION OF FRONT AREAS

(*a*) $\theta < 45°$.

Let A_F be the absorption factor calculated for the front area only. A_F.

When $\theta < 45°$, $A_F = \int_{x=0}^{x=2} \frac{ds}{dx} e^{-\mu r x} \, dx$,

where s is given by equation (16*a*).



$\theta = 45°$

Figure 6.

Writing equation (16) in the form α_r

$$s = \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3 + \ldots,$$

we have

$$\frac{ds}{dx} = \alpha_1 + 2\alpha_2 x + 3\alpha_3 x^2 + \ldots.$$

On integration by parts and suitable rearrangement,

$$A_F = \frac{\alpha_1}{\mu r}(1 - e^{-2\mu r}) + \frac{2!\,\alpha_2}{\mu^2 r^2}\{1 - e^{-2\mu r}(1 + 2\mu r)\} + \frac{3!\,\alpha_3}{\mu^3 r^3}\left\{1 - e^{-2\mu r}\left(1 + 2\mu r + \frac{4\mu^2 r^2}{2!}\right)\right\} + \ldots\ldots.$$

Put ρ

$$2\mu r = \rho.$$

Then

$$A_F=\frac{2\alpha_1}{\rho}(1-e^{-\rho})+\frac{2^2\,2!\,\alpha_2}{\rho^2}\{1-e^{-\rho}(1+\rho)\}+\frac{2^3\,3!\,\alpha_3}{\rho^3}\left\{1-e^{-\rho}\left(1+\rho+\frac{\rho^2}{2!}\right)\right\}$$
$$+\ldots+\frac{2^n\,n!\,\alpha_n}{\rho^n}\left(1-e^{-\rho}\sum_{k=0}^{n-1}\frac{\rho^k}{k!}\right)+\ldots \quad\ldots\ldots(18).$$

$1-e^{-\rho}\sum_{k=0}^{n-1}\frac{\rho^k}{k!}$ may be written $e^{-\rho}\left(e^{\rho}-\sum_{k=0}^{n-1}\frac{\rho^k}{k!}\right)$,

or by expansion of e^{ρ}, $\qquad e^{-\rho}\sum_{k=n}^{\infty}\frac{\rho^k}{k!}$.

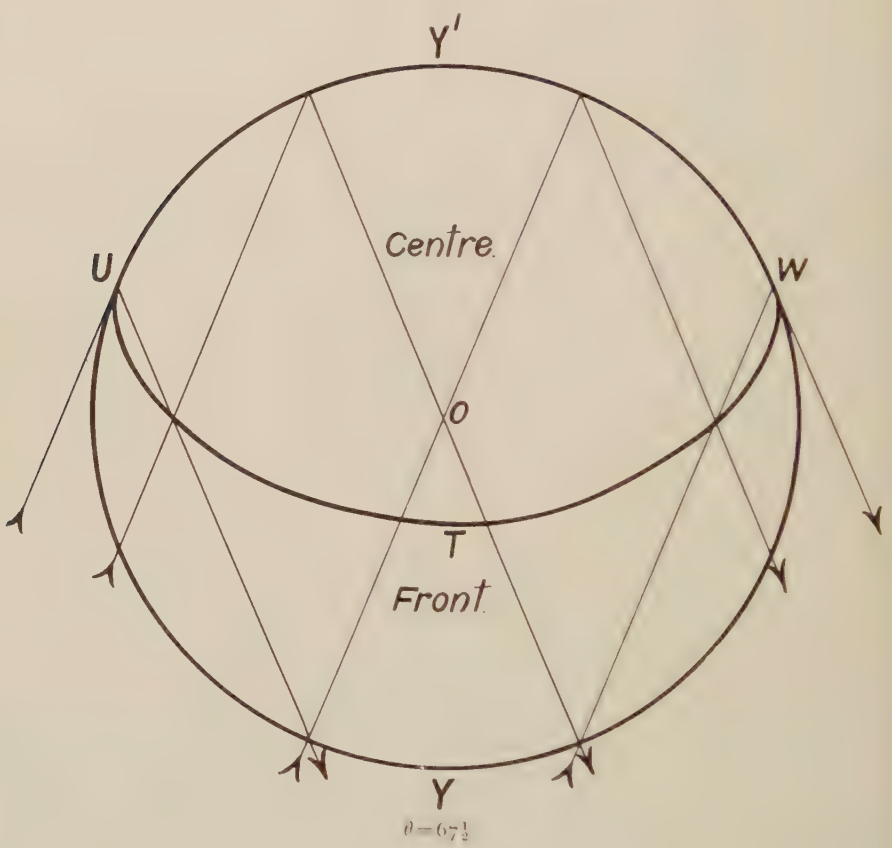


Figure 7.

It is convenient to write equation (18) thus:

$$\sum_{n=1}^{7}\frac{2^n\,n!\,\alpha_n}{\rho^n}\left(1-e^{-\rho}\sum_{k=0}^{n-1}\frac{\rho^k}{k!}\right)+e^{-\rho}\sum_{n=9}^{\infty}\frac{2^n\,n!\,\alpha_n}{\rho^n}\sum_{k=n}^{\infty}\frac{\rho^k}{k!} \quad\ldots\ldots(19).$$

R The order of summation may be reversed in the second or remainder term *R*, which becomes

$$e^{-\rho}\sum_{l=0}^{\infty}\sum_{n=9}^{\infty}\frac{2^n\,n!\,\alpha_n\,\rho^l}{(n+l)!}.$$

This lies between $e^{-\rho}\sum_{n=9}^{\infty} 2^n\alpha_n$ and $e^{-\rho}\sum_{l=0}^{\infty}\sum_{n=9}^{\infty}\frac{2^n\,9!\,\alpha_n\rho^l}{(9+l)!}$.

A convenient approximation is

$$R=e^{-\rho}\sum_{n=9}^{\infty}2^n\alpha_n\left\{\frac{1}{2}+\frac{1}{2}\sum_{l=0}^{\infty}\frac{9!}{(9+l)!}\rho^l\right\} \qquad \text{.....(20).}$$

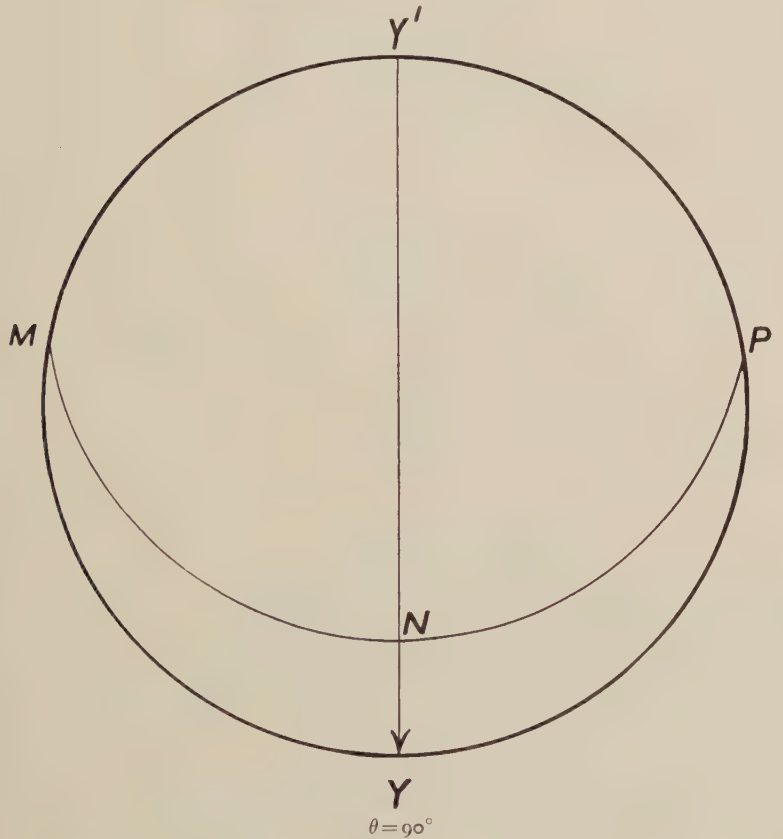


Figure 8.

The value of A_F is calculated from equation (18) by substitution for α_1, α_2, etc., from equation (16*a*) up to the term containing α_7. Since α_4, α_6, etc., are zero, this involves five terms only. The series either converges so rapidly that no further terms are required (when $\mu r>5$), or so slowly that it would be tedious to calculate sufficient terms (when $\mu r<5$). In the latter case an estimate is made of the remainder after the term containing α_7, using equation (20).

(*b*) $\theta>45°$.

When $\theta>45°$ $$A_F=\int_{x=0}^{x=2\sin 2\theta}\frac{ds}{dx}\,e^{-\mu r x}\,dx,$$

where s is given by equation (16*a*).

On putting $2\mu r \sin 2\theta = \rho$, we find that equation (18) still holds good but 2^n is replaced by $(2 \sin 2\theta)^n$. The calculations are with this exception the same as for $\theta < 45°$.

§ 5. THE ABSORPTION FACTOR—CORRECTIONS FOR BACK AND CENTRE AREAS

(*a*) *Back areas:* $0 < \theta < 30°$.

A_B

$$A_B = \int_{x=4\sin\theta}^{x=2} \frac{ds}{dx} e^{-\mu r x}\, dx,$$

where s is given by equation (16*b*).

As before, $$s = \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3$$

and $$\frac{ds}{dx} = \alpha_1 + 2\alpha_2 x + 3\alpha_3 x^2 + \ldots .$$

Integrating by parts and rearranging, we have

$$A_B = \frac{\alpha_1}{\mu r}(e^{-4\mu r \sin\theta} - e^{-2\mu r}) + \frac{2!\,\alpha_2}{\mu^2 r^2}\{e^{-4\mu r \sin\theta}(1 + 4\mu r \sin\theta) - e^{-2\mu r}(1 + 2\mu r)\} + \ldots .$$

ρ, τ Put $$2\mu r = \rho,\ 4\mu r \sin\theta = \tau.$$

Then

$$A_B = \alpha_1\left(\frac{4\sin\theta}{\tau}e^{-\tau} - \frac{2}{\rho}e^{-\rho}\right) + 2!\,\alpha_2\left\{\frac{4^2\sin^2\theta}{\tau^2}e^{-\tau}(1+\tau) - \frac{2^2}{\rho^2}e^{-\rho}(1+\rho)\right\}$$

$$+ \ldots + \alpha_n n!\left(\frac{4^n \sin^n\theta}{\tau^n}e^{-\tau}\sum_{k=0}^{n-1}\frac{\tau^k}{k!} - \frac{2^n}{\rho^n}e^{-\rho}\sum_{k=0}^{n-1}\frac{\rho^k}{k!}\right) + \ldots \quad \ldots\ldots(21).$$

The value of A_B for angles 0 to 30° is calculated from equation (21) by substitution for α_1, α_2, etc., from equation (16*b*) up to the term containing x^7. Successive terms get more and more equal to those for the front area obtained by using equation (18), so that the remainder may easily be estimated from equation (20) when a slight allowance based on the lower terms has been made.

(*b*) *Centre areas:* $\theta < 30°$.

The area not included in the previous calculations is grouped together under the heading "centre area". The total area πr^2 consists of front F, back B and centre C. When $\theta < 30°$, the front and back areas are given by putting $x = 2$ into equation (16). On subtraction of the values of F and B so found from πr^2 the central area is found to be

$$C = \frac{4r^2\cos^2\theta}{\sin\theta}\int_{q=0}^{q=\tan\theta}\sqrt{\{1/(1+q^2) - \cos^2\theta\}}\, dq.$$

The length of path in this area varies between $x = 2$ and $x = 2\sec\theta$. As an approximation put

$$x = (1 + \sec\theta).$$

A_C The contribution A_C of the central area is then given by

$$A_C = \frac{4}{\pi}\frac{\cos^2\theta}{\sin\theta}\int_{q=0}^{q=\tan\theta}\sqrt{\{1/(1+q^2)\cos^2\theta\}}\, dq\, e^{-\mu r(1+\sec\theta)} \quad \ldots\ldots(22).$$

This elliptic integral has already been evaluated in series form, as part of equation (16).

(*c*) *Centre areas:* $30° < \theta < 90°$.

Equation (22) is suitably modified to allow for the difference in area and length of path. As θ approaches 90° this procedure becomes unsatisfactory, owing to the increasing variation in length of path throughout the area. Fortunately it is not necessary to make any calculations for angles just below 90° as the absorption factors may be found by interpolation. When $\theta = 90°$ a special exact formula, which replaces this estimate, is applicable. It is also possible from a modification of the formula used for the front area to make an independent estimate of the correction for the centre area at angles above 45°.

(*d*) $\theta = 90°$.

At 90° the former centre area includes the whole specimen. The absorption factor is deduced from equation (18) but instead of $2\mu r = \rho$, we have $4\mu r = \rho$ and 4^n replaces 2^n, while α_1, α_2, etc., are deduced from equation (17) by an expansion similar to equation (16).

§6. THE ABSORPTION FACTOR: SPECIAL CASES

In general the absorption factor is found by adding together the contributions of the front, back and centre areas, thus:

$$A = A_F + A_B + A_C.$$

A_F is calculated as in equation (18), A_B as in (21) and A_C either as in (22), or by a modification of this formula. A_B is only operative when $0° < \theta < 30°$. For the special case of $\theta = 0°$, equation (21) becomes identical with equation (18), so that we have

$$\text{when } \theta = 0°, \quad A = 2A_F.$$

For the special case where $\theta = 90°$, equation (18) is sufficient, but the constants α_1, α_2, etc., are obtained from equation (17) and not equation (16).

When $\mu r > 5$ it is found that except when $\theta = 0°$ and $\theta = 90°$, A_B and A_C may be neglected in comparison with A_F. Further, the remainder term in equation (19) is negligible. In the first term of equation (19) the factor $\left(1 - e^{-\rho} \sum_{k=0}^{n-1} \frac{\rho^k}{k!}\right)$ is approximately equal to unity.

This is not the case for the higher terms. Hence it is important to conclude the series at $n = 7$, thus:

when $\mu r > 5$

$$A = \sum_{n=1}^{7} \frac{\alpha_n n!}{\mu^n r^n} \qquad \text{.....(23),}$$

where $\alpha_1, \alpha_2 \ldots \alpha_7$ are given as coefficients of $x, x^2 \ldots x^7$ in equation (16*a*).

Equation (23) differs from Claassen's equation (4) because it is a finite series, and therefore the question of convergency does not arise. Equation (23) does not hold for $\theta = 0°$ and $\theta = 90°$.

For extremely large μr only the first term of equation (23) need be considered. From equation (16*a*)

$$\alpha_1 = \frac{1}{\pi}\left\{1 - \frac{\log(\sec\theta + \tan\theta)}{\sec\theta\tan\theta}\right\}.$$

Hence *for extremely large* μr

$$A = \frac{1}{\pi\mu r}\left\{1 - \frac{\log(\sec\theta + \tan\theta)}{\sec\theta\tan\theta}\right\} \qquad \ldots\ldots(24).$$

Equation (24) is virtually equivalent to Rusterholz's formula for the case of great absorption. According to Rusterholz, the expression in brackets is

$$\left\{1 + \frac{\cos^2\theta}{2\sin\theta}\log\frac{\cos 2\theta + \sin\theta}{(1+\sin\theta)(1+2\sin\theta)}\right\}.$$

This can of course be simplified to the form given in equation (24).

When $\mu r > 10$ three terms of equation (23) are sufficient for 1 per cent accuracy.

$$A = \frac{1}{\pi\mu r}\left\{1 - \frac{\log(\sec\theta + \tan\theta)}{\sec\theta\tan\theta}\right\} + \frac{\sin 2\theta}{2\pi\mu^2 r^2} + \frac{1}{\pi\mu^3 r^3} \times\left[-\frac{1}{4} + \frac{3}{4}\cos^2\theta\left\{\frac{1}{2} + \frac{1}{2}\frac{\log(\sec\theta + \tan\theta)}{\sec\theta\tan\theta}\right\}\right] \qquad \ldots\ldots(25).$$

At 0° this expression must be doubled, and for angles just above 0 a certain correction is required, to allow for the contribution from the back of the specimen. This contribution shows itself on the film as a doubling of the line and only occurs at extremely low angles. At 90° equation (25) is not applicable, the correct value being, however, only slightly different from that given by this equation. The true formula is as follows:

$$A_{(\theta=90°)} = \frac{1}{\pi\mu r} - \frac{1}{16\pi\mu^3 r^3} \qquad \ldots\ldots(26).$$

The important point to realize is the steady increase in the complexity of the formulae as μr decreases. When $5 < \mu r < 10$ it is necessary to add two terms to equation (25), but the calculations are still not very complicated. When $\mu r < 5$ complications immediately ensue at low angles. It is then necessary to bring in corrections for the higher terms of the series and for the contributions of the centre and back of the specimen. Finally, these corrections become so important that the method of calculation becomes less reliable. When $\mu r < 2$ the Claassen graphical

Table 2. Values of absorption coefficient

	$\theta = 0°$		$\theta = 22\frac{1}{2}°$		$\theta = 45°$		$\theta = 67\frac{1}{2}°$		$\theta = 90°$
μr	Graphical	Calculated	Graphical	Calculated	Graphical	Calculated	Graphical	Calculated	Assumed
5	4·73	4·34	18·40	18·71	48·4	48·0	80·1	81·0	100
4	7·67	7·17	21·30	21·56	51·0	50·8	81·4	82·3	100
3	14·25	13·66	26·75	27·35	56·0	55·3	83·5	84·4	100
2	30·10	30·10	40·50	40·50	64·1	64·1	87·8	88·0	100
1·5	45·60	45·70	53·10	53·10	71·6	71·6	91·0	90·3	100
1	66·80	66·60	71·10	70·60	82·4	81·4	94·3	94·0	100

method supersedes the calculations, and the graphical method can be employed successfully down to $\mu r = 0$. Between $\mu r = 1$ and $\mu r = 5$ it is possible to compare the results obtained by the two methods. It will be seen from table 2 that the agreement is perfect when $\mu r = 2$. The calculated values are more correct when $\mu r > 2$ and the graphical values when $\mu r < 2$.

§ 7. TABLES OF ABSORPTION FACTORS

The two tables of absorption factors were calculated from the formulae given in the present paper except for the range below $\mu r = 2$, where Claassen's graphical method was used. The table given in Claassen's paper (table 1, p. 59), which formed the basis of the calculations for this method, was checked and in some instances corrected. The values when $\theta = 0°$ especially required emendation. With these changes it has been shown that the graphical method gives exactly the same results as the present formulae for the case where $\mu r = 2$ so that there is some reason to suppose that the calculations are accurate for still smaller values of μr, for which the graphical method is even more reliable.

The values when $\mu r < 5$ are absolute values of A. When $\mu r > 5$ the values are expressed relative to the value for 90°. If it is desired to find the absolute value, it is only necessary to calculate the value of A when $\theta = 90°$ from equation (26), which is applicable for all values of μr above 5. The tables are used by plotting the values of A (absolute or relative), as given, against the appropriate value of $\sin^2\theta$. It will be found that the curve obtained is so nearly linear that it can be easily and accurately drawn. For values of θ intermediate between those given in the tables the absorption values can easily be estimated by interpolation. An example is given of the curve for $\mu r = 5$.

The values given in these tables are directly applicable to the lines belonging to the layer of zero order in rotation photographs, if the specimen is cylindrical about the axis of rotation. To apply them to powder photographs some modification is required if the specimen is diluted. In place of μr the following substitution is made:

$$\mu r = \frac{\mu}{\rho} \cdot \frac{m}{\pi r h},$$

where μ/ρ is the mass absorption coefficient and m is the mass of powder in a specimen of height h and radius r.

It is assumed that the particles are uniformly distributed in the specimen and that each particle is so small that absorption inside it is negligible. If the diluent has no appreciable absorption coefficient it is only necessary to know μ/ρ, m, r and h. The following technique has been successfully employed by the author. The specimen is diluted with Canada balsam, which also serves as adhesive. It is then mounted on a hair in the form of a cylinder about 0·4 mm. in diameter. After the exposure has been made, the specimen is removed from the powder camera and the average cross-section in the portion exposed to X rays is measured by means of a microscope. It is then placed on a weighed microscope cover slip and the length h is measured, after any irregular portions at the top and bottom have been cut off

Table 3. Absorption factors 100 A when $\mu r < 5$

$\sin^2\theta$	0	0·1464	0·3290	0·5000	0·6710	0·8536	1·00
θ	0°	22½°	35°	45°	55°	67½°	90°
μr							
0·0	100	100	100	100	100	100	100
0·1	84·70	84·80	—	84·9	—	85·0	85·1
0·2	71·20	71·60	—	71·9	—	72·4	72·9
0·3	60·00	60·60	—	61·4	—	62·7	63·5
0·4	51·00	51·70	—	52·7	—	54·5	55·6
0·5	43·50	44·20	—	45·8	—	47·8	49·0
0·6	36·90	37·80	—	39·8	—	42·3	43·6
0·7	31·40	32·40	—	34·8	—	37·8	39·3
0·8	26·80	27·85	—	30·5	—	33·7	35·6
0·9	23·00	24·10	—	27·1	—	30·5	32·4
1·0	19·77	20·95	—	24·2	—	27·85	29·5
1·1	16·98	18·28	—	21·70	—	25·50	27·15
1·2	14·59	16·00	—	19·54	—	23·50	25·10
1·3	12·56	14·03	—	17·70	—	21·70	23·35
1·4	10·84	12·33	—	16·11	—	20·10	21·80
1·5	9·38	10·91	—	14·69	—	18·66	20·50
1·6	8·11	9·73	—	13·52	—	17·46	19·32
1·7	7·10	8·71	—	12·47	—	16·41	18·24
1·8	6·15	7·80	—	11·54	—	15·42	17·30
1·9	5·37	7·02	—	10·74	—	14·59	16·44
2·0	4·71	6·35	—	10·05	—	13·84	15·67
2·1	4·16	5·79	—	9·44	—	13·15	14·93
2·2	3·67	5·31	—	8·89	—	12·50	14·26
2·3	3·24	4·86	—	8·38	—	11·89	13·65
2·4	2·865	4·47	—	7·91	—	11·35	13·09
2·5	2·55	4·12	—	7·50	—	10·86	12·56
2·6	2·27	3·82	—	7·11	—	10·40	12·11
2·7	2·02	3·55	—	6·75	—	9·98	11·67
2·8	1·803	3·30	—	6·41	—	9·62	11·27
2·9	1·607	3·08	—	6·10	—	9·25	10·89
3·0	1·436	2·885	4·43	5·82	7·24	8·89	10·54
3·1	1·288	2·705	4·22	5·58	6·97	8·57	10·21
3·2	1·159	2·55	4·02	5·35	6·72	8·30	9·90
3·3	1·049	2·415	3·84	5·14	6·48	8·04	9·61
3·4	0·955	2·29	3·68	4·95	6·26	7·78	9·33
3·5	0·871	2·17	3·525	4·77	6·05	7·55	9·06
3·6	0·796	2·06	3·385	4·60	5·85	7·33	8·81
3·7	0·729	1·968	3·255	4·44	5·66	7·11	8·58
3·8	0·670	1·875	3·14	4·29	5·49	6·92	8·36
3·9	0·617	1·787	3·03	4·15	5·33	6·73	8·15
4·0	0·568	1·706	2·925	4·02	5·17	6·53	7·94
4·1	0·525	1·629	2·825	3·89	5·02	6·35	7·74
4·2	0·488	1·563	2·73	3·77	4·88	6·18	7·55
4·3	0·453	1·500	2·645	3·66	4·75	6·02	7·38
4·4	0·420	1·445	2·555	3·56	4·62	5·87	7·21
4·5	0·391	1·390	2·485	3·47	4·50	5·73	7·05
4·6	0·364	1·343	2·41	3·38	4·39	5·61	6·89
4·7	0·340	1·300	2·34	3·29	4·28	5·47	6·75
4·8	0·316	1·259	2·275	3·21	4·18	5·35	6·61
4·9	0·2945	1·222	2·21	3·13	4·08	5·25	6·47
5·0	0·2755	1·189	2·155	3·05	3·99	5·14	6·35

Table 4. Relative values of A for large values of μr

$\sin^2\theta$	0	0·1464	0·3290	0·5000	0·6710	0·8536	1·00
θ / μr	0°	22½°	35°	45°	55°	67½°	90°
5·0	4·34	18·71	33·93	48·00	62·8	81·0	100
5·5	3·50	17·78	32·95	47·05	62·0	80·4	100
6·0	2·91	17·01	32·10	46·25	61·3	79·9	100
6·5	2·44	16·40	31·40	45·60	60·7	79·5	100
7·0	2·12	15·89	30·80	45·05	60·2	79·2	100
7·5	1·83	15·46	30·30	44·50	59·8	78·9	100
8·0	1·61	15·09	29·85	44·05	59·4	78·6	100
9·0	1·26	14·45	29·10	43·35	58·7	78·1	100
10·0	1·02	13·98	28·55	42·75	58·2	77·7	100
11·0	0·84	13·59	28·10	42·25	57·8	77·4	100
12·0	0·69	13·26	27·70	41·85	57·4	77·2	100
13·0	0·59	13·00	27·40	41·50	57·1	77·0	100
14·0	0·50	12·78	27·15	41·25	56·9	76·8	100
15·0	0·44	12·59	26·90	41·00	56·7	76·7	100
20·0	0·25	11·93	26·05	40·20	55·9	76·2	100
25·0	0·16	11·55	25·60	39·70	55·4	75·8	100
30·0	0·11	11·29	25·25	39·35	55·1	75·6	100
40·0	0·06	10·97	24·85	38·95	54·8	75·3	100
50·0	0·04	10·79	24·60	38·70	54·5	75·1	100
75·0	0·02	10·54	24·30	38·40	54·2	74·9	100
100	0·01	10·42	24·10	38·20	54·0	74·8	100
200	0·00	10·18	23·90	38·00	53·8	74·6	100
∞	0·00	10·07	23·65	37·70	53·6	74·4	100

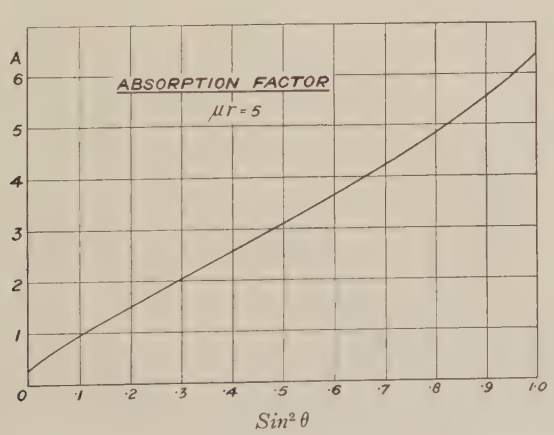


Figure 9.

with a razor blade. The Canada balsam is dissolved away by means of toluene, which is afterwards soaked up with blotting paper. The hair can then be removed and finally only the powder is left on the cover slip, which is reweighed, the weight of powder being obtained by difference. If necessary, where the absorption coefficient of the powder is small a correction may be made for absorption in the Canada balsam. This proceeds as for a powder containing ingredients of different absorption coefficients.

Suppose that the powder contains two ingredients, a mass xm of an ingredient with mass absorption coefficient $\mu_1\ \rho_1$ and a mass $(1-x)\ m$ of a second ingredient with mass absorption coefficient $\mu_2\ \rho_2$. The value of μr is then given by

$$\mu r = \left[\left(\frac{\mu_1}{\rho_1}\ xm\right) + \left\{\frac{\mu_2}{\rho_2}\ (1-x)\ m\right\}\right] \frac{1}{\pi r h}.$$

Evidently only the value of x is required in addition to the previous data. It must be emphasized that the particles of the two ingredients need to be perfectly mixed, so that no portion of the specimen contains enough of one material to cause an abnormal absorption.

Finally, a word of warning is necessary with regard to the utility of the absorption factor. The observed intensities are attributed to a number of factors. Some depend on the structure of the crystals while others depend on the angle of reflection, changing steadily as the angle passes from small to large values. Amongst the factors concerned must be included the temperature factor, which causes a steady decrease in the intensities as θ increases. This is the opposite of the absorption factor, which reduces the intensities at low angles much more than those at high angles. In many instances the effects of the absorption factor and the temperature factor almost cancel one another. Consequently a comparison of observed intensities with calculated intensities will often show better agreement if no allowance is made for the absorption correction, unless at the same time the values are corrected for the temperature effect.

The absorption factor is calculated on the assumption that the whole specimen is bathed in a homogeneous parallel beam of X rays. Actually there is no such thing as a parallel beam of X rays. Satisfactory results are obtained, however, if the beam diverges from a distant point. Provided the radius of the specimen is small in comparison with this distance the validity of the calculations is not seriously affected. Trouble will certainly arise if narrow slits are introduced between the source of X rays and the specimen, for they have the effect of producing variations in the intensity-distribution at different points of the specimen. In other words, the specimen must be *completely* bathed in the incident beam.

Slight irregularities in the shape of the specimen are to a large extent compensated by continuous rotation of it during the exposure. It is essential to use a specimen containing a large number of particles in order to obtain a sufficiently uniform distribution of scattering matter at all angles. If the powder is packed in Lindemann glass tubing, this must be so thin-walled that absorption in the glass is negligible.

§ 8. ACKNOWLEDGMENT

The author thanks Professor W. L. Bragg, F.R.S. for his kind interest in the work, which was carried out in the Physical Laboratories of the University of Manchester.

REFERENCES

(1) L. W. McKeehan. *J. Frankl. Inst.* **193**, 231 (1922).
(2) G. Greenwood. *Phil. Mag.* (7), **3**, 963 (1927).
(3) H. Möller and A. Reis. *Z. phys. Chem. A.* **139**, 425 (1928).
(4) A. Claassen. *Phil. Mag.* (7), **9**, 57 (1930).
(5) A. Rusterholz. *Helv. Physica Acta*, **4**, 68 (1931).

538.313

THE MAGNETIC FORCE AT A POINT ON ITS AXIS DUE TO A CURRENT IN A HELICAL COIL OF ONE TURN

By G. F. C. SEARLE, Sc.D., F.R.S.,
University Lecturer in Experimental Physics, Cambridge

Received June 1, 1935

ABSTRACT. The incomplete circuit formed by a single turn ABC of a helix of pitch p, which lies on a cylinder of radius a, is completed by a straight link CA, of length p, parallel to the axis of the cylinder. Through any point O on this axis are taken rectangular axes Ox, Oy, Oz. Of these, Ox is the axis of the cylinder, and Oz is parallel to the shortest distance between Ox and CA. The components H_1, H_2, H_3 of the magnetic force at O, due to a current in the circuit, are found for the case when p/a is small.

§ 1. MAGNETIC FORCE AT ORIGIN

Let the helix $ABPC$, figure 1, lie on a cylinder of radius a; let it be right-handed, with pitch p. The circuit is completed by a portion CDA, of length p, of a generator of the cylinder. The axis of the helix is the axis of x, and right-handed axes are used. We will take as origin O the point on the axis of x at which

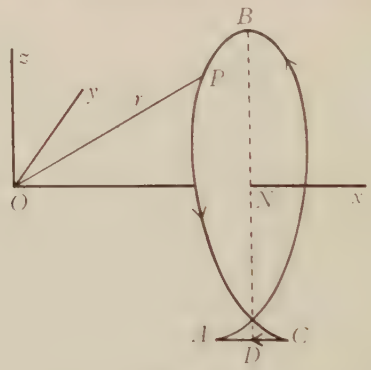


Figure 1.

the magnetic force is to be found. The mean plane of the helix, parallel to Oyz, cuts the helix in B and the axis in N, where $x=h$, and bisects CDA in D. The point P lies on the helix, the plane POx has turned in the positive direction through angle θ from the plane zOx, and $OP=r$. The link CDA is in the plane zOx. If the coordinates of P be ξ, η, ζ, and if $p/2\pi=q$, we have

$$\xi=h+p\theta/2\pi=h+q\theta, \quad \eta=-a\sin\theta, \quad \zeta=a\cos\theta.$$

Then

$$d\xi/d\theta=q, \quad d\eta/d\theta=-a\cos\theta, \quad d\zeta/d\theta=-a\sin\theta.$$

If ds be an element of arc, $ds/d\theta=\sqrt{(q^2+a^2)}=b$. If λ, μ, ν be the direction cosines of the forward direction of ds,

$$\lambda=q/b, \quad \mu=-a\cos\theta/b, \quad \nu=-a\sin\theta/b.$$

Let dH_1, dH_2, dH_3 be the components, along Ox, Oy, Oz, of the magnetic force at O due to the element ds, or $bd\theta$, at P. Then

$$dH_1=ibd\theta r^{-3}(\nu\eta-\mu\zeta)=ia^2d\theta r^{-3},$$
$$dH_2=ibd\theta r^{-3}(\lambda\zeta-\nu\xi)=iad\theta r^{-3}\{h\sin\theta+q(\cos\theta+\theta\sin\theta)\},$$
$$dH_3=ibd\theta r^{-3}(\mu\xi-\lambda\eta)=iad\theta r^{-3}\{-h\cos\theta+q(\sin\theta-\theta\cos\theta)\}.$$

We have $$r^2=a^2+(h+q\theta)^2.$$

If $q\theta$ be small compared with l, where $l^2=a^2+h^2$, we have

$$r^{-3}=l^{-3}\{1+c_1q\theta+c_2q^2\theta^2+c_3q^3\theta^3+\ldots\},$$

where $$c_1=-3h/l^2, \quad c_2=-\tfrac{3}{2}(a^2-4h^2)/l^4, \quad c_3=\tfrac{5}{2}h(3a^2-4h^2)/l^6.$$

For H_1 we have, as far as q^3,

$$H_1=ia^2\int_{-\pi}^{\pi}r^{-3}d\theta=\frac{ia^2}{l^3}\int_{-\pi}^{\pi}(1+c_1q\theta+\ldots)\,d\theta=\frac{2\pi ia^2}{l^3}(1+\tfrac{1}{3}c_2\pi^2q^2)$$
$$=\frac{2\pi ia^2}{l^3}\left\{1-\frac{\pi^2q^2(a^2-4h^2)}{2l^4}\right\} \qquad \ldots\ldots(1).$$

The axial component H_1 of the magnetic force equals, accurately, that due to a uniform cylindrical current sheet, of radius a and strength i/p, extending from the plane $x=h-\frac{1}{2}p$ to the plane $x=h+\frac{1}{2}p$. Accurately

$$H_1=\frac{2\pi i}{p}\left\{\frac{h+\frac{1}{2}p}{\{(h+\frac{1}{2}p)^2+a^2\}^{\frac{1}{2}}}-\frac{h-\frac{1}{2}p}{\{(h-\frac{1}{2}p)^2+a^2\}^{\frac{1}{2}}}\right\}.$$

When we expand as far as p^2, we obtain (1).

We next consider H_2. As far as q^3,

$$\{h\sin\theta+q(\cos\theta+\theta\sin\theta)\}\,r^{-3}=\{f_0+f_1q+f_2q^2+f_3q^3\}\,l^{-3},$$

where $$f_0=h\sin\theta, \quad f_1=\cos\theta+\theta\sin\theta\,(1+c_1h),$$
$$f_2=c_1\theta\cos\theta+\theta^2\sin\theta\,(c_1+c_2h), \quad f_3=c_2\theta^2\cos\theta+\theta^3\sin\theta\,(c_2+c_3h).$$

When the following integrations extend from $-\pi$ to π, we have

$$\int\sin\theta d\theta=0, \quad \int\cos\theta d\theta=0, \quad \int\theta\sin\theta d\theta=2\pi, \quad \int\theta\cos\theta d\theta=0,$$
$$\int\theta^2\sin\theta d\theta=0, \quad \int\theta^2\cos\theta d\theta=-4\pi, \quad \int\theta^3\sin\theta d\theta=2\pi^3-12\pi, \quad \int\theta^3\cos\theta d\theta=0.$$

If H_2' be the part of H_2 which is due to the helix ABC,

$$H_2'=\frac{ia}{l^3}\int_{-\pi}^{\pi}(f_0+f_1q+\ldots)\,d\theta$$
$$=\frac{2\pi iaq}{l^3}\left[1-\frac{3h^2}{l^2}-\frac{q^2}{2l^2}\left\{3(\pi^2-8)\frac{a^2-4h^2}{l^2}-5(\pi^2-6)\frac{h^2(3a^2-4h^2)}{l^4}\right\}\right].$$

The current along CA from C to A produces at O magnetic force H_2'' along Oy. We have

$$H_2''=ia\int_{-\frac{1}{2}p}^{\frac{1}{2}p}\frac{dt}{\{a^2+(h+t)^2\}^{\frac{3}{2}}}=\frac{ia}{l^3}\int_{-\pi q}^{\pi q}(1+c_1t+c_2t^2+\ldots)\,dt$$

$$=\frac{2\pi iaq}{l^3}\left\{1-\frac{\pi^2q^2(a^2-4h^2)}{2l^4}\right\}.$$

If H_2 be the total magnetic force at O along Oy, $H_2=H_2'+H_2''$, and

$$H_2=\frac{2\pi iaq}{l^3}\left\{\frac{2a^2-h^2}{l^2}-\frac{(\pi^2-6)\,q^2}{2l^6}(4a^4-27a^2h^2+4h^4)\right\} \qquad \ldots\ldots(2).$$

We now find H_3. As far as q^3, we have

$$\{-h\cos\theta+q(\sin\theta-\theta\cos\theta)\}\,r^{-3}=\{g_0+g_1q+g_2q^2+g_3q^3\}\,l^{-3},$$

where $g_0=-h\cos\theta$, $g_1=\sin\theta-\theta\cos\theta\,(1+c_1h)$,

$g_2=c_1\theta\sin\theta-\theta^2\cos\theta\,(c_1+c_2h)$, $g_3=c_2\theta^2\sin\theta-\theta^3\cos\theta\,(c_2+c_3h)$.

On integration from $-\pi$ to π, we have

$$\int g_0d\theta=0,\quad \int g_1d\theta=0,\quad \int g_3d\theta=0$$

$$\int g_2d\theta=2\pi(3c_1+2c_2h)=-6\pi h(4a^2-h^2)\,l^{-4}.$$

Hence

$$H_3=\frac{ia}{l^3}\int_{-\pi}^{\pi}(g_0+g_1q+\ldots)\,d\theta$$

$$=-\frac{6\pi iaq^2h}{l^7}(4a^2-h^2) \qquad \ldots\ldots(3).$$

§ 2. MAGNETIC FORCE NEAR ORIGIN

If we wish to find the components of the magnetic force at a point Q, of coordinates x, y, z, near O, we must, in place of the value of dH_1 used in § 1, write

$$dH_1=ibd\theta\,r_Q^{-3}\{\nu(\eta-y)-\mu(\zeta-z)\},$$

where $r_Q^2=(\xi-x)^2+\ldots=r^2\{1-2(\xi x+\eta y+\zeta z)/r^2+(\xi x^2+\ldots)/r^2\}$.

Similar expressions hold for dH_2, dH_3. We first expand r_Q^{-3} in powers of x, y, z and of r^{-1} and then expand r^{-3}, r^{-5}, ... in powers of $q\theta$. We then insert the values of ξ, η, ζ, λ, μ, ν in terms of θ and integrate with respect to θ as in § 1.

§ 3. A PAIR OF HELICES

If the helix be turned about its axis through 180°, the signs of H_2 and H_3 are reversed. Hence we obtain the following accurate result:

Let equal currents flow in the same direction in two equal coaxial helices ABC, $A'B'C'$, each of one turn, with their mean planes coincident. If the angle between the two planes containing (1) the axis and the link AC, (2) the axis and the link $A'C'$ be π, the magnetic force at any point on the axis has no component at right angles to the axis.

§ 4. COMPUTATION BY MAGNETIC SHELL

It is, perhaps, surprising that the leading term in H_2 should be of the *first* order in q. We can, however, verify the result. The helix ABC (figure 1) and the link CA form the edge of a magnetic shell of strength i. We take the shell to have a cylindrical part and a plane part. The cylindrical part has Ox for axis and extends from the helix in the positive direction of x. The plane part, which closes the positive end of the cylindrical part, is a disk of radius a with Ox for axis. The part of the shell on the positive side of the plane $x=h+\frac{1}{2}p$ is symmetrical about Ox, and the magnetic force at O due to it has no component along Oy. The remainder of the shell is a strip of width k, where

$$k=\tfrac{1}{2}p\,(1-\theta/\pi)=q\,(\pi-\theta).$$

With the direction of the current employed in § 1, the magnetism on the outer side of the shell is positive. If da be the thickness of the shell and if σ be the surface density of magnetism on the outer surface, $\sigma da=i$.

The magnetic force at O, due to a pole m at a point ξ, η, ζ on the cylinder, has a y-component

$$-\frac{m\eta}{(a^2+\xi^2)^{\frac{3}{2}}}=\frac{ma\sin\theta}{r^3}.$$

If a be increased to $a+da$, the y-component due to m becomes

$$m\sin\theta\left\{\frac{a}{r^3}+\frac{d}{da}\left(\frac{a}{r^3}\right)da\right\}.$$

Hence, if dH_2 be the y-component of the magnetic force at O due to an element of shell of area dS at ξ, η, ζ, we have

$$dH_2=\sigma dS\sin\theta\,\frac{d}{da}\left(\frac{a}{r^3}\right)da=idS\sin\theta\,\frac{\xi^2-2a^2}{r^5}.$$

If we do not go beyond the first power of q, we may put $\xi^2=h^2$, $r^2=l^2=a^2+h^2$, and $dS=kad\theta$. Then, to the first power of q,

$$H_2=\frac{iaq\,(h^2-2a^2)}{l^5}\int_{-\pi}^{\pi}(\pi-\theta)\sin\theta d\theta=\frac{2\pi iaq\,(2a^2-h^2)}{l^5},$$

as found in § 1.

§ 5. ACKNOWLEDGMENT

Mr K. T. Li, Emmanuel College, Cambridge, B.Sc., National Central University, Nankin, China, has kindly verified the analysis in this paper.

537.322.2:546.57

LONGITUDINAL THERMOELECTRIC EFFECT: (5) SILVER

By J. L. CH'EN, M.Sc. (Yenching)

AND

WILLIAM BAND, M.Sc., Professor of Physics,
Yenching University, Peiping, China

Received April 10, 1935

ABSTRACT. Improvements in technique are described by means of which the experimental results become more consistent than before, and the analysis becomes simpler to carry out. The homogeneous thermoelectric constants for silver are found, and their dependence upon longitudinal tension is shown. A critical temperature near 200° C. for zero tension is found at which the thermoelectric constants suffer an abrupt change. Reasons are given for supposing that there is an intimate connection between this transition and the ordinary elastic limit of the wire.

§ 1. IMPROVEMENTS IN TECHNIQUE

The general method and idea of this work can be learned from the previous papers under the present general title [1, 2, 3]. The thermoelectric e.m.f. in this case was measured with the Paschen galvanometer previously described in part 3, but the thermal regulator was similar to that described in part 2. The chief improvement over the method used in part 3 was to surround the heating-system with a cylindrical metal jacket containing a water-cooled cavity: this cylinder was closed at both ends except for small apertures at the axes for the specimen and thermocouples, and it maintained the outside conditions for the heating system at a definite known temperature throughout.

To improve the probe method of investigating the temperature-distribution, the probe elements were led into the axial hole from opposite ends of the axis and joined in the middle to form the thermojunction. Although the two elements have not quite the same thermal properties, yet this device, when made to slide to and fro along the axis in exploring the temperature-distribution, certainly creates much less disturbance of the temperatures than did the previous probe entering entirely from one end. As has previously been pointed out, the probe, while for obvious reasons not as precise as could be wished, is the only practicable means available for the purpose in view. The mechanical and thermal maltreatment of the specimen involved in the welding of a thermo-junction on to or into it, most decidedly has to be avoided if any significance is to be attached to the e.m.fs. as due to a homogeneous

metal. However, to check errors arising in the positions of the temperatures given by the travelling junction, it is necessary to have one or more junctions fixed in position along the axis; these also suffer the same temperature error due to their superficial contact with the specimen, but it is assumed that the axial hole has an almost uniform temperature-distribution in planes perpendicular to the axis.

These fixed thermocouples also served to check the constancy of the temperatures at fixed points while the travelling couple was moving through the system during the mapping process.

In the previous work the analysis was carried out by trial and error; in the present work a graphical device eliminates the necessity for this. The integrals of the temperature-distribution in the equation:

$$E = a\int G.dT + b\int G^2.dT \qquad a, b$$

are found as before for each temperature-distribution, and then the straight lines are plotted against the axes a and b as the two variables. The common point of intersection of the whole set of lines gives, theoretically, the required values of a and b. Actually, the weighted centre of the smallest area containing all these lines will in practice be the point accepted. A visual picture of the precision of the result is thus very nicely obtained.

§ 2. RESULTS

In figure 1 are shown seven typical temperature-distributions split into straight-line sections for analysis. For each temperature-distribution the specimen was carried through the set of tensions investigated, and the resulting e.m.fs. were

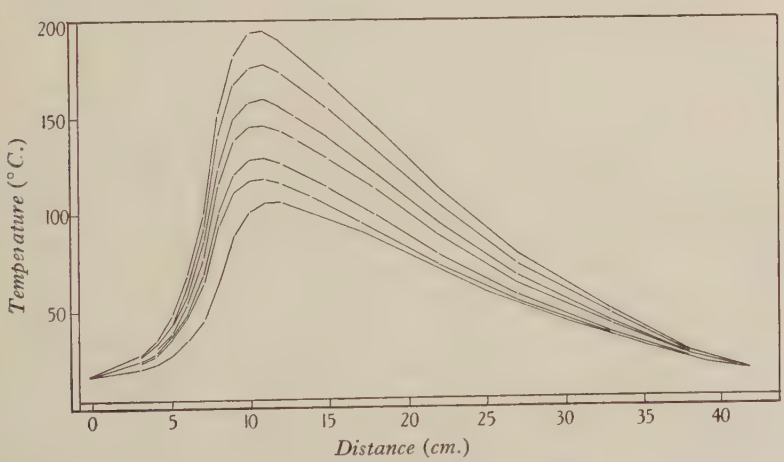


Figure 1. Typical temperature-maps.

recorded. It was early noted that after a certain tension had been passed the changes in e.m.f. ceased to be reversible and a permanent alteration of the {e.m.f., tension} curves took place. The wire was therefore replaced and the observations were repeated, care being taken always to keep the tension below this critical value.

In figure 2 is shown the graphical analysis of the data for zero tension on the first specimen. The common point of intersection of all the lines is well defined, indicating a satisfactory precision for the resulting values of the thermoelectric constants. The values indicated are

$$a = -6{\cdot}1 \times 10^{-10},$$

and

$$b = \quad 8{\cdot}6 \times 10^{-12}.$$

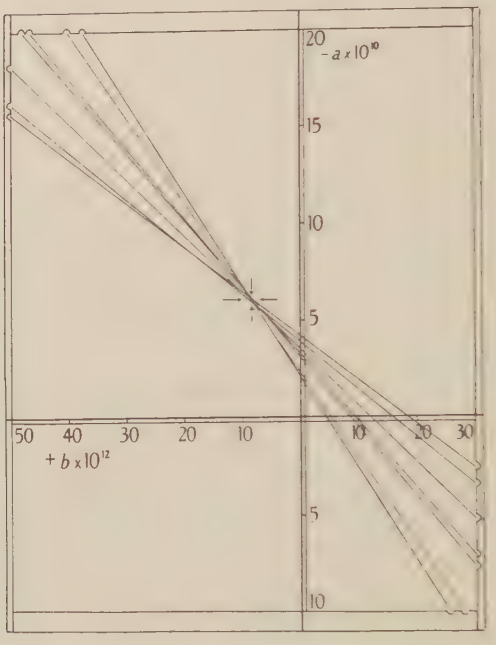


Figure 2. Graphical analysis for thermoelectric constants a and b.

In table 1 are given the e.m.fs. in microvolts for three temperature-distributions for the second specimen. It is to be noticed that the e.m.fs. drop quickly from a maximum initial value at zero tension to approximately a saturation value near 300 kg./cm.^{2}

Table 1

Curve	S (kg./cm.^{2})							
	0	100	200	300	400	500	600	700
1	0·166	0·129	0·105	0·080	0·080	0·080	0·080	0·080
2	0·270	0·221	0·180	0·173	0·169	0·166	0·170	0·172
3	0·295	0·229	0·230	0·225	0·241	0·252	0·247	0·252

The values obtained for zero tension are, in this case

$$a = -\ 6{\cdot}2 \,.\, 10^{-10}\ \text{V.-cm./(° C.)}^2$$

$$b = \quad 14{\cdot}5 \,.\, 10^{-12}\ \text{V.-cm.}^2\text{/(° C.)}^3$$

In figure 3 are given the values of the constants as functions of tension. The curves are drawn from the empirical formulae:

$$-a.10^{10} = 3{\cdot}4 + 2{\cdot}8 \times \exp(-0{\cdot}00625.S),$$
$$b.10^{12} = 5{\cdot}0 + 9{\cdot}5 \times \exp(-0{\cdot}00369.S).$$

The asymptotes, unlike those found in the case of copper[1], appear to be practically parallel with the S axis as in normal saturation curves.

It will be noticed that the two specimens gave almost identical values for the constant a but that the two values for b are quite different. It was observed that the general behaviour of the e.m.f. with changing tension was the same for the two specimens. Since both specimens were taken from the same piece of wire it is not

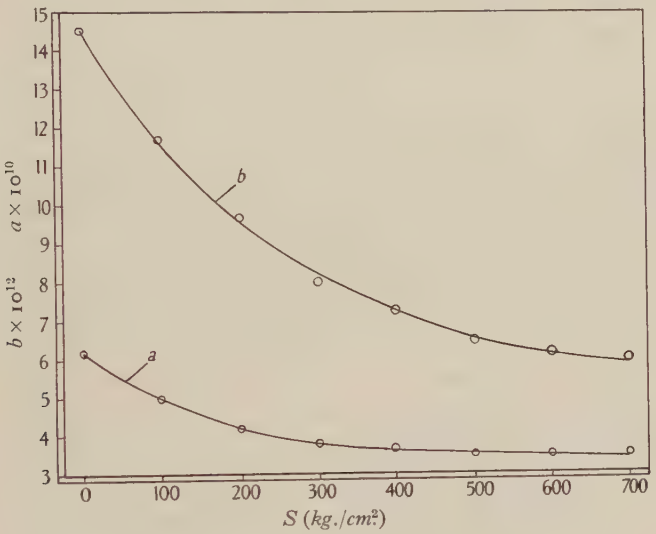


Figure 3. Dependence of a and b upon longitudinal strain.

possible to explain the difference in b as due to a difference in the metal or its impurities. However, in view of the experience gained with the first specimen, the second specimen was not heated to so high a maximum temperature. It may be that there is a rapid variation of b with temperature as we pass from the range of temperatures used in the second set to the higher temperatures of the first set of readings. But in that case it is difficult to see how the analysis can give definite results on the assumption of a constant value for b throughout the temperatures existing in the specimen (from about 20° C. up to the maximum used, say 150° C.). Slight twisting of the wire may account for the change in b, which must in that case be very sensitive to the twist, for care was taken not to twist the wire during mounting in the apparatus.

These effects require further detailed investigation of the dependence of the e.m.f. upon temperature-range, and upon state of torsion of the material.

The afore-mentioned transition was next investigated, use being made of several different specimens in turn as each was carried over the transition temperature.

The specimen was heated to a temperature as near under the transition as possible without actually passing it, the normality being checked by carrying the specimen through a tension cycle. The process was repeated at a slightly higher maximum

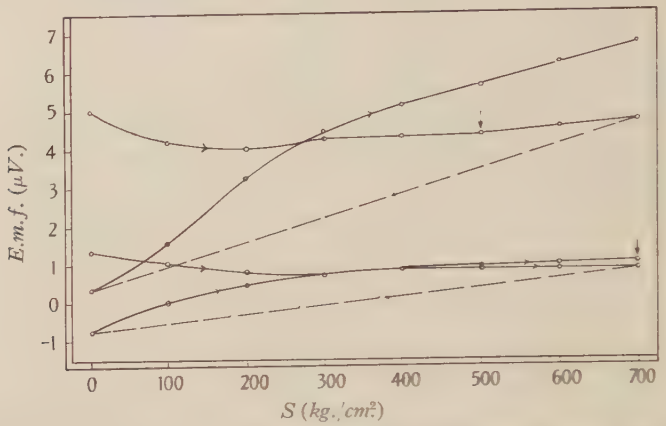


Figure 4. Discontinuity produced at elastic limit in e.m.f., tension curves. The arrows indicate the tensions S at which the hysteresis first appeared in the two specimens.

temperature (at the peak of the distribution curve) and an entirely different {tension, e.m.f.} curve resulted; see figure 4. The initial heating was carried out while the wire was kept at the maximum tension to be used in the cycle. It was

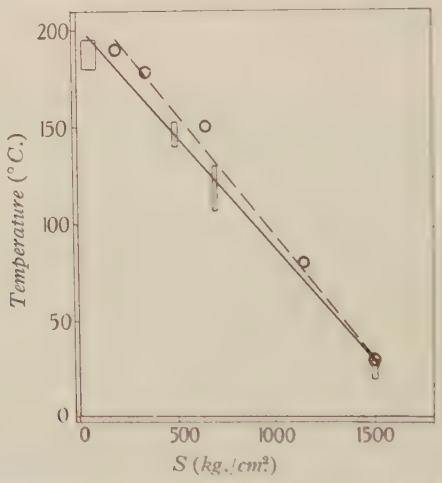


Figure 5. Thermoelectric discontinuity and the elastic limit. The full line and rectangles show the thermoelectric data, the dotted line and circles the mechanical data.

found, first, that if this maximum tension was increased, the maximum temperature required to produce the transition was not so high. If the wire was stretched beyond its elastic limit at room-temperature its behaviour corresponded with the phenomena found after the change at higher temperatures and smaller tensions. Figure 5 shows how the transition temperature depended upon the maximum

tension. It seemed most probable that this curve followed the course of ordinary elastic limit as it changes with temperature. A direct test of this suggestion gave the dotted curve in the same figure. Evidently the correct explanation is that the part of the wire heated above the critical temperature is stretched beyond its elastic limit and so behaves abnormally, either in that it forms a thermocouple with the normal wire, or in that its stretching allows bodily motion of the rest of the wire through the temperature-field.

§ 3. NOTE ON CONVENTIONS AS TO SIGNS

In the *International Critical Tables* the sign convention for the Thomson coefficient is that when the electrons have positive specific heat the effect is positive. This convention is modified as follows, in the present work, to include the new e.m.f.: If a temperature-gradient tends to cause a spontaneous flow of electrons down the gradient from hot to cold, the effect is to be taken as positive. This convention includes the above convention for the Thomson coefficient, and also enables a definite sign to be given to the two homogeneous coefficients studied in the present work. Thus if, as in the effect for silver, the spontaneous positive current flows out of the specimen at the end where the steeper temperature-gradient terminates, the electrons must be flowing up the steeper gradient and down the flatter gradient; the effect is then negative. With this convention, the formula used for analysis in previous papers on this subject, has been written

$$\mathbf{F} = t\mathbf{G} + a\,(\mathbf{G}.\mathbf{G})^{\frac{1}{2}}\,\mathbf{G} + b\,(\mathbf{G}.\mathbf{G})\,\mathbf{G},$$

because when the coefficients are positive the positive current, supposed to be driven by the field $\mathbf{F}$, is in the same sense as the temperature-gradient $\mathbf{G}$, the temperature increasing with positive displacement. The electrons flow down the temperature-gradient, or opposite to the direction of $\mathbf{G}$.

REFERENCES

(1) P. C. Feng and W. Band. *Phys. Soc. Proc.* **46**, 515 (1934).
(2) T. H. Pi and W. Band. Page 852 of this volume.
(3) M. K. Li and W. Band. Page 859 of this volume.

538.66:546.74

THERMOMAGNETIC HYSTERESIS IN NICKEL WIRE

By Y. K. HSÜ, M.S. (Yenching)

AND

WILLIAM BAND, M.Sc., Professor of Physics,
Yenching University, Peiping, China

Received April 10, 1935

ABSTRACT. The thermomagnetic e.m.f. in a nickel wire of diameter 1·0 mm. is produced by longitudinal magnetic fields and by longitudinal stretching in one-half of each of a set of various temperature-distributions. The hysteresis is examined as dependent both upon magnetism and upon tension. It is found that for zero tension the results agree with those of Broili[1] to a first approximation, although there is still a measurable increase in e.m.f. not detected by Broili when the maximum temperature is raised above the Curie point. This is interpreted as due to the so-called Benedick e.m.f. depending on the temperature-gradient. The Curie-point increase is still greater in the results for stretched wire. The form of the hysteresis curve also becomes anomalous in that the e.m.f. reverses sign, relative to the original value, in two quadrants of the loop. The change in Thomson potential-gradient produced by various magnetizing fields, at different tensions, is shown graphically as a function of temperature, to a first approximation obtained by neglecting the Benedick e.m.f.

§ 1. INTRODUCTION

The thermomagnetic properties of nickel wire have been studied by Heinz Broili[1]. The specimen was centrally heated to a uniform temperature while the ends were maintained at a constant lower temperature. The part of the wire in the temperature-drop at one end was subjected to a magnetic field, while the part of the wire in the other temperature-drop was outside the magnetic field. The electromotive force produced in the wire by this arrangement was studied and found to show positive hysteresis. The e.m.f. was also found to vary with temperature-difference, but when the temperature-gradient was changed without change of the temperature-difference, no detectable change in the e.m.f. occurred.

Work of somewhat similar nature was done by Chang and Band[2] on iron wire. Here the hysteresis was found to be negative, and a qualitative theory was proposed to explain this result. This particular theory, however, would demand a negative hysteresis even in nickel, contrary to Broili's results. More work was therefore planned on nickel wire to check this point.

In the work of Pi and Band[3] nickel was first examined for the Benedick effect which according to Broili's results either did not exist or was masked by the larger modification of the Thomson e.m.f. in the magnetic field. The work gave positive

results; the e.m.f. was however of a smaller order than that observed by Broili, the maximum change produced by a magnetizing field of about 200 gauss being only a few microvolts. In this work all the e.m.fs. were considered as being due respectively to the magnetic modification of the Thomson potential-gradient and of the two new thermoelectric constants a and b, which were associated with the square and the cube of the temperature-gradient[4].

In the work reported in the present paper, nickel* has again been examined in an arrangement essentially similar to that of Broili. A uniform magnetic field was superposed on a part of the wire also suffering a temperature-gradient, while the parts of the wire which had varying magnetization were kept approximately at a uniform temperature. In addition to this, the wire could be stretched on one side to a different tension from that on the other, by means of a mechanical clamp electrically insulated from the specimen. Thus the pure thermomechanical e.m.f. caused by the modification of Thomson potential-gradient under longitudinal strain could be investigated. Also, as a further improvement over Broili's work, not only could the temperature-difference be obtained by a fixed thermocouple, but the whole temperature-distribution could be mapped throughout the system by means of thermocouple probes made of platinum against platinum-rhodium alloy. This mapping permits an estimate of the small e.m.fs. due to the Benedick effect previously ignored. Since our values of e.m.f. are precise to 0·5 μV. we find it quite obvious that the Benedick effect does produce detectable values of e.m.f. If we want to get the pure modification of the Thomson gradient alone, this Benedick effect should be calculated and accounted for, at least approximately.

This is particularly marked in the phenomena near the Curie point. According to Broili's results, any increase in the maximum temperature after it has already passed the Curie point (near 350° C.) is accompanied by no further increase in thermomagnetic e.m.f. This indicated that the Thomson potential-gradient is no longer modified by the magnetic field in those parts of the specimen which lie above the Curie-point temperature. In the present work the same is true, but only approximately. Under zero tension, for instance, as the maximum temperature is increased from 320° C. to 380° C. there is already an increase of 1·0 μV., while a further increase of temperature to 500° C. produced a further e.m.f. of about 5·0 μV. The results when the tension was 8 kg. were much more marked; the change from 320 to 380 being accompanied by an increase in e.m.f. of 4·5 μV. We might explain these results as due to a shift of Curie point with tension, but the known dependence is in the opposite direction[5] and much smaller than that required for the present problem. We prefer to explain the e.m.f. as due to the temperature-gradient changing as the maximum temperature is raised. The Benedick e.m.f. on the two sides of the temperature map—one e.m.f. occurring in wire at zero tension and zero magnetic field, and the other in magnetized strained wire—will both be changed by increase in the temperature-gradient, and not in general by the same amount. Judged from the results of Pi and Band already mentioned, the effect is of the right order.

* Pure nickel wire manufactured by Schering-Kahlbaum AG.

§ 2. PROCEDURE AND APPARATUS

Figure 1 shows the arrangement. The test wire was stretched between two spring dynamometers, the clamps being simple screws with rubber washers bearing on the test wire. Microscope micrometers at either end, focused on two marks on the wire, permitted measurement of the elongation of the specimen and a check on its stability of position. The set-up was mounted on a heavy wooden stand which supported also the two water-cooling tubes and the heating-system symmetrically on either side of the central clamp. This clamp was rigidly made of brass casting with milled beds for the jaws, which were of snug fit; this eliminated all detectable shift in their position when they were subject to horizontal forces. The jaws were 0·5 cm. in width along the direction of the wire, from which they were electrically

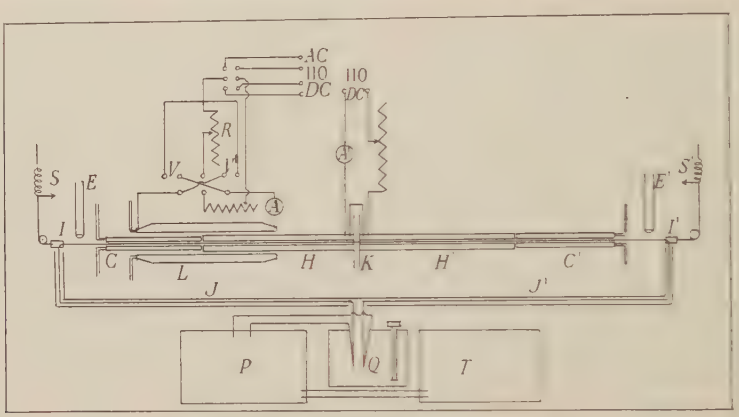


Figure 1. Arrangement of apparatus.

A, *A'*, ammeters. *C*, *C'*, water-cooling tubes. *E*, *E'*, microscope micrometers. *H*, *H'*, electrical heaters. *I*, *I'*, insulated clamps. *J*, *J'*, shielding tube round specimen. *T*, thermostat for standard cell. *V*, *V'*, special switching device for hysteresis loop. *K*, central clamp. *L*, water-cooled solenoid. *P*, K-type potentiometer. *Q*, oil bath with stirrer. *R*, resistance control to stopping point in hysteresis loop. *S*, *S'*, spring dynamometers.

insulated by mica sheets. The slate heaters abutted immediately upon both sides of the clamp jaws. They were non-inductively wound with eureka wire to the same total resistance and were connected in series. An actual test of the eureka wire gave its composition as copper 55 per cent, nickel 40 per cent and zinc 5 per cent. This was the least magnetic resistance-wire available; a check on its permeability indicated even less magnetism than would be expected from the percentage of nickel present in it, but it is admitted that this may be a source of error, particularly in regard to the form of the hysteresis loops obtained, and in partially disturbing the uniformity of the magnetic field about the specimen.

The temperature-distributions produced by different working currents were approximately mapped by thermocouple probes used in preference to couples annealed to the test wire for reasons thoroughly discussed elsewhere[6]. The element platinum and an alloy, platinum (90 per cent) + rhodium (10 per cent), gave a linear calibration curve to above the maximum temperature to be observed. One such

couple was kept at a fixed position in the centre of each heater to check the constancy of the temperature-distribution during both the mapping and the subsequent thermomagnetic work. During the thermomagnetic observations continual control of the heating-current to a constant value was employed to ensure constancy of the temperature-distribution, the probes being removed from the apparatus after the mapping. Four temperature-distributions actually used are shown in figure 2.

The magnetizing solenoid of the Moullevigen form was so placed that the uniform field was superposed on that part of the wire which was subjected to the greatest temperature-gradient; it extended from the minimum temperature in the water-cooled part to the point of maximum temperature.

The magnetizing cycle was controlled by a switching-device made after the design of A. W. Smith[7], which carried the current through a complete cycle and

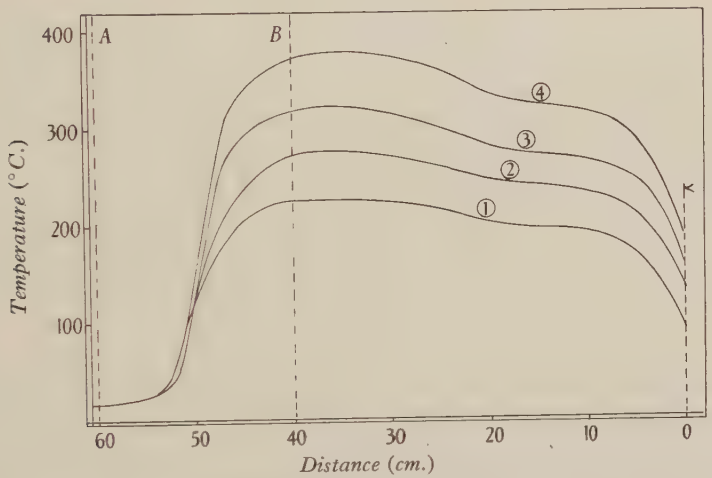


Figure 2. Temperature-distributions. The numbers correspond to those referred to in later diagrams of thermomagnetic hysteresis, etc. *A–B* is the extent of the uniform field.

stopped it at any desired point ranging from the maximum (10 A.) to as small a value as we pleased; actually 0·1 A. was the minimum used for the hysteresis work. This device is of fundamental usefulness, permitting as it does complete elimination of errors due to zero-wander and departure from the cyclic state of the specimen during measurement of the hysteresis loop. The loops actually obtained were perfectly symmetrical, never departing from symmetry by more than the 0·5 μV., which was the minimum difference detectable with certainty on the potentiometer used. By a few rapid oscillations of the switch between individual readings the specimen was maintained in the cyclic state as long as was required.

The specimen wire extended from the tension-clamps at either end to an oil bath kept at uniform temperature by gentle stirring. Copper-wire connections to the K-type potentiometer completed the essentials of the electrical circuit.

In spite of all care in the construction of the apparatus, the temperature-distributions actually obtained were not perfectly symmetrical, although they were nearly

so. Usually it was found that even when no tension and no magnetic field are applied, a small but definite e.m.f. was measurable on the potentiometer, presumably on account of the lack of symmetry. This e.m.f. was usually of the order of 2 or 3 μV., and was observed to vary in magnitude so long as the heating-currents were not kept exactly constant. It was found possible, however, by constant watch over the heating-system, to maintain this initial e.m.f. constant after the steady state had been reached. Since it had proved quite sensitive to slight changes in the temperature-distribution, it was therefore assumed that the temperature had remained constant throughout a set of observations if it was reproduced after the end of the set. If the initial e.m.f. varied by more than 0·5 μV. during a set of observations in spite of constant watch over the heating-system, then the set was rejected. Only seldom did this occur after the equipment had been brought into proper working order.

From the very beginning doubt was felt about the effect of the central clamp on the homogeneity of the specimen and the subsequent e.m.fs. to be studied. Runs were therefore taken of e.m.f. against magnetic field with the clamp both in action and entirely removed. No real difference was found. Runs were then taken of e.m.f. against tension on one side of the clamp and were followed by similar runs with tension on the other side. Here it was observed that slight differences occurred which may have been due to either of two factors: (*a*) the clamp may have introduced asymmetrical pressures at the clamped part which were not reliably duplicated when the tension changed in direction; or (*b*) since the temperature-distribution was known to be not quite symmetrical, the Benedick e.m.f. would change differently on the two sides. We believe that the last-mentioned is the real cause, chiefly because after returning to zero tension on both sides the initial e.m.f. was satisfactorily reproduced even though the clamp had not yet been released.

Finally, hysteresis runs were taken of e.m.f. against magnetizing field at different tensions, both with and without the use of the central clamp. In almost all cases the results were less consistent when the clamp was not in use; but otherwise there was no real difference between the two sets of results. The lack of mutual consistency among the results when the clamp was released must be due to slight shifts of the specimen permitted by the somewhat non-rigid clamps at the dynamometers. It was therefore decided that the use of the central clamp was justified for the purposes of (i) investigating the e.m.f. produced by a difference of tension in the two parts of the specimen and (ii) keeping the wire fixed mechanically during other observations in which tension was required in the part of the wire suffering magnetization.

The chief drawback to the use of the central clamp was that it necessarily caused a considerable dip, which could not be measured very precisely, in the temperature curves. This point must be considered when the distribution is analysed for the Benedick correction terms.

§ 3. RESULTS

Thermomechanical e.m.f. and hysteresis. The results shown in figure 3 were obtained by keeping the magnetic field zero and the tension on one side of the central

clamp zero. The e.m.f. is plotted as a function of the tension on the other side; it is to be regarded essentially as due to a modification of the homogeneous thermo-electric constants in that part of the wire which is subjected to the longitudinal strain, which thus destroys the balance of the complete circuit. The total e.m.f. obtained was of the order of 20 μV., most of which must be due to modification of the Thomson potential-gradient. However, from results not yet published we estimate that several microvolts will be due to modification of the other constants a and b mentioned earlier. In particular there is no true saturation of the e.m.f. as the maximum temperature passes the Curie point, and the increase is not negligible in comparison with the total e.m.f.

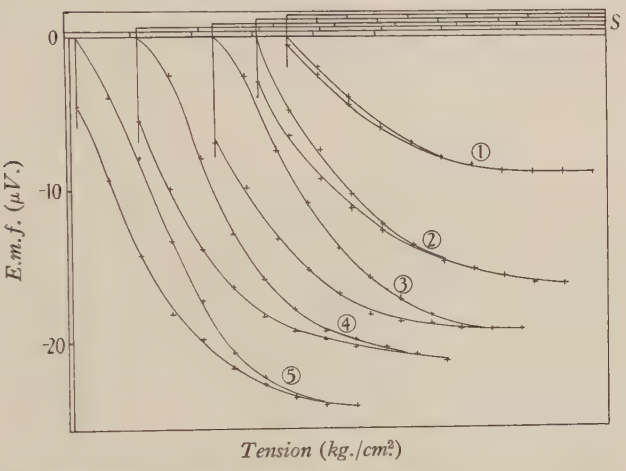


Figure 3. Thermomechanical e.m.f. and hysteresis. The intervals on the horizontal S axis represent* tension-increases of 2 kg. in every case; the hysteresis loops are drawn with their origins displaced for the sake of clarity. Curve (5), as in succeeding diagrams, corresponds with a temperature-distribution (not shown in figure 2) whose maximum was about 500° C.

The hysteresis observed is normal in form, and seems to increase with increasing temperature—a circumstance which is probably of some significance for the explanation of the hysteresis. If the latter were simply elastic lag, it would be expected to increase with rising temperature.

Thermomagnetic e.m.f. and hysteresis. The initial rise of e.m.f. against increasing magnetizing field is shown for the different temperature-distributions studied, in figure 4, for zero tension (and for a tension of 8 kg. for comparison later). Their lack of saturation as the temperature passes the Curie point has already been mentioned; otherwise they are saturation curves very similar in form to the known curve for the square magnetization plotted against field for nickel.

The hysteresis found for zero tension is shown in figure 5. There is very little difference between the various curves at different temperature-distributions. They are also similar to the curves obtained by Broili referred to previously. It is to be noted however that at no point in the cycle could the e.m.f. actually be brought back

* A tension of 1 kilogram corresponds to 127·3 kg./cm².

to the initial zero value, until of course after the final demagnetization. For the highest three temperatures the curves might be extrapolated to intersect near the zero axis, but the actual experiment failed to reveal such a point; the minimum

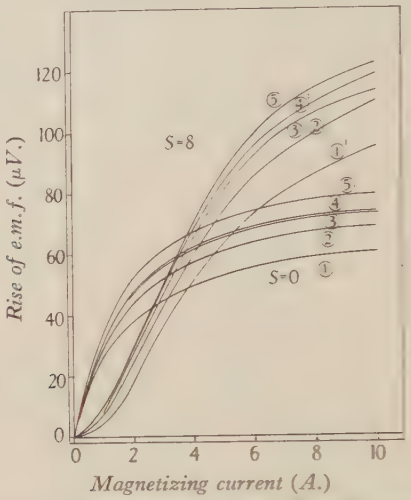


Figure 4. Thermomagnetic e.m.f. Initial rise of e.m.f. against magnetizing field H, for the five temperature-distributions indicated for tensions of 0 and 8 kg. respectively.

e.m.f. obtainable at this inverse peak of the cycle was always about 2 μV. This may have been due to a really non-vanishing minimum, or it may have been due to a zero minimum that was too sharp to be detected.

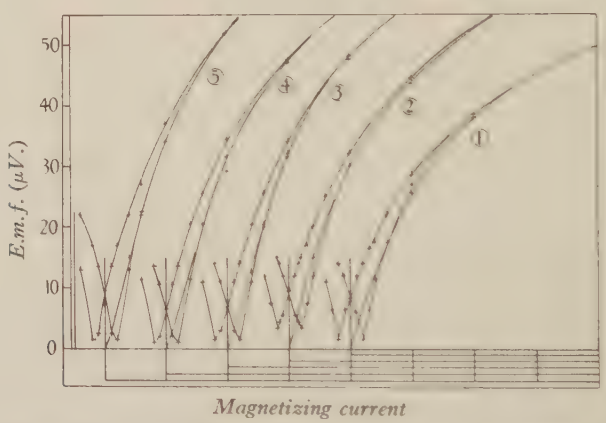


Figure 5. Thermomagnetic hysteresis loops. The intervals on the horizontal H axis represent 2 A. in every case*. The loops have been drawn with their origins displaced horizontally for clarity.

Mechanical-magnetic equivalence. From the above it will be seen that the e.m.f. produced by longitudinal strain is opposite in direction to that due to longitudinal magnetization. This is in harmony with the known fact that longitudinal strain produces radial magnetization in nickel. It was thought worth while to take a few

* A magnetizing field of 10 amperes corresponds to 303 gauss.

uns of tension against magnetic field so balanced that the e.m.f. produced should emain at the initial zero. The results are given in figure 6. It is possible to assume hat the two disturbances, because they annul one another, are essentially equivalent n magnitude, and this may furnish us with a good method of finding quantitatively he magnetizing effect of longitudinal strain. It is pointed out, however, that there s probably a skin effect in both the longitudinal and the transverse magnetizations: he former is already well-known, and the latter must exist to prevent infinite con-entration of polarity down the axis of the wire. If the two skin-depths are almost equal, no magnetization will be produced when the tension and magnetic field are balanced in the above manner, and they can be truly considered as equivalent. If he skin-depths are different no such conclusion can be drawn.

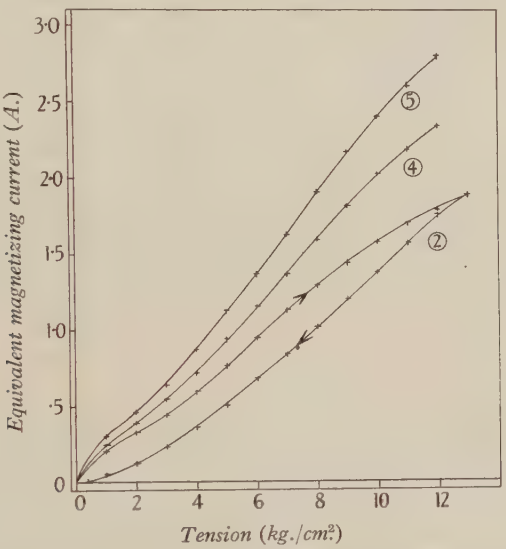


Figure 6. Mechanical-magnetic equivalence. Curves of magnetizing field H against tension S balanced for zero thermo-e.m.f. A complete hysteresis curve is shown in the case of lowest temperature-distribution; the numbers refer to temperatures previously mentioned.

The hysteresis curve shown in the same figure is of a form which evidently indicates that the radial magnetization follows the tension-change with less hysteresis than that with which the longitudinal magnetization follows the magnetizing field*.

Effect of strain on thermomagnetic hysteresis. Not only is the magnitude of the e.m.f. produced by magnetization changed when the wire is subjected to tension, but also the form of the hysteresis loop is greatly modified. This phenomenon is shown in figure 7. The variation of initial e.m.f. under zero field appearing in this diagram is of no significance; it arises merely from the fact that only half of the wire is stretched, the half not in the magnetizing solenoid being free from tension. The dotted curve passing through the initial points is simply the pure thermomechanical e.m.f. for rising tension already shown in figure 3. Had the other half of the wire also been stretched to the same tensions in each case this e.m.f. would have been

* See additional note on demagnetization phenomena.

eliminated, but obviously no change would have thereby been made in the form of the hysteresis loops. As regards this hysteresis itself it will be noticed that the area

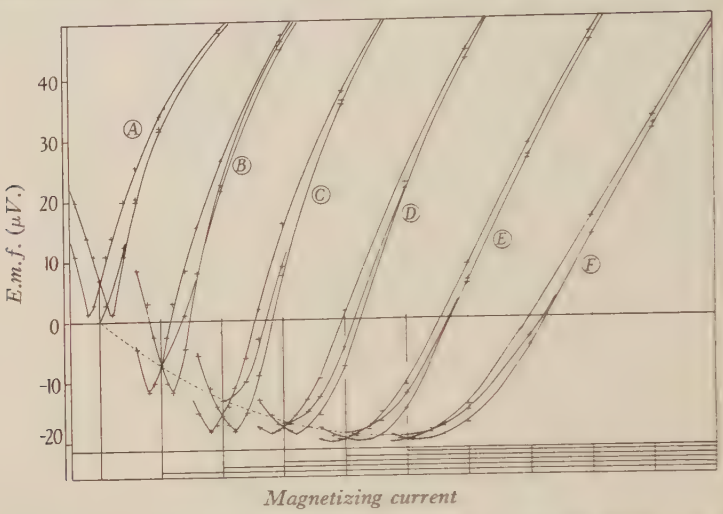


Figure 7. Thermomagnetic hysteresis and tension. Intervals on the horizontal H axis are in steps of 2 A. A refers to zero tension; B to 2 kg.; C to 4 kg.; D to 6 kg.; E to 8 kg.; and F to 10 kg. Only the lower part of one-half of each cycle is shown here to permit the use of a sufficiently large scale. Temperature-distribution (3) obtains in each case.

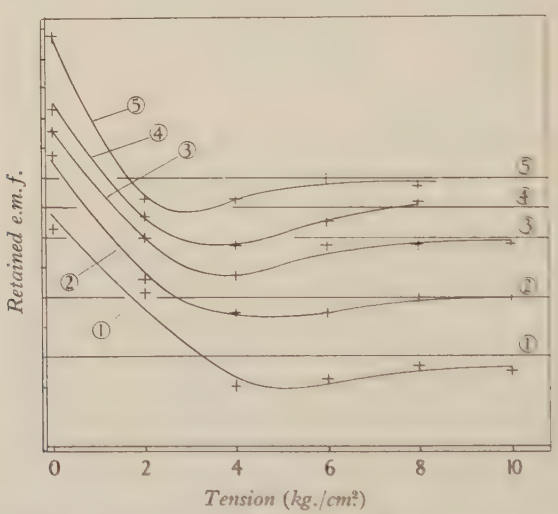


Figure 8. Retentivity of thermomagnetic e.m.f. The values of the retained e.m.f. relative to the initial e.m.f. at the beginning of the magnetizing cycle are shown for the four temperature-distributions indicated. The vertical axis shows intervals of 2μV. and the zero axis for each curve is shown by the correspondingly numbered horizontal line.

of the loop seems to increase a little for greater tensions, but the most obvious change is in the general shape of the curves. For quite small tensions the point of minimum e.m.f. in the hysteresis becomes greatly negative, but the minimum flattens out for greater tensions. In figure 8 are shown curves of retentivity against tension for the

different temperature-distributions; in every case the retentivity changes from the normal positive anomalous negative values for small tensions returning asymptotically towards zero for high tensions.

§ 4. INTERPRETATION OF RESULTS

To a first approximation only, the electromotive forces observed in the present work can be considered as due to modifications of the Thomson potential-gradient in that part of the wire which is subject to the disturbances considered, namely longitudinal magnetization and longitudinal strain, or a combination of both. This approximation, as has already been emphasized, ignores the homogeneous e.m.f. associated with the square and cube of the temperature-gradient. The present results are quite precise enough to proceed with the second approximation in which these new e.m.fs. are included. This analysis is reserved for a later paper, pending completion of an exhaustive study of the constants a and b* on the lines begun in the work of Pi and Band already quoted. *a, b*

To a first approximation, therefore, all the curves here obtained are really curves of Thomson potential-gradient against magnetizing field or tension as the case may be, except that the initial undisturbed value remains unknown or the position of the true origin of the curves is not determined.

Now the Thomson e.m.f. E produced in a wire between two points A and B will be given by *E*

$$E=\int_A^B \sigma dT. \qquad \sigma$$

If the wire is in two sections AB and BC in which σ has different values σ_1 and σ_2, and if the temperature-distribution is symmetrical with respect to the point B, then

$$E=\int_A^B \sigma_1 dT+\int_B^C \sigma_2 dT$$
$$=\int_A^B (\sigma_1-\sigma_2)\, dT.$$

In general σ will be a function of both H, the magnetizing field, and S the tension. *H, S*
Also, if we write $\sigma\,(HS)$ to indicate this functional relation, we shall have

$$-\sigma\,(H_1S_1)+\sigma\,(H_2S_2)=\Delta\sigma,$$

also a function of temperature. In other words, the modification of the Thomson potential-gradient produced by a given ΔH or by a given ΔS will depend considerably upon the temperature T. Thus in setting the e.m.f. E, since *T*

$$E=\int_A^B \{\sigma\,(H_2S_2)-\sigma\,(H_1S_1)\}\, dT$$
$$=\int_A^B \Delta\sigma . dT,$$

where S_1 and H_1 are usually both zero in our particular work, we cannot assume that E will be directly proportional to the temperature-difference T_B-T_A. However,

* See page 905.

since $\Delta\sigma$ is supposed to be a function only of temperature, and not of position in the wire (if the magnetization and tension are uniform), we can consider E to be a function only of $(T_B - T_A)$, and independent of the temperature-gradient in the wire.

In particular if the temperature T_B is increased to T_B', say, a shorter part of the wire now exhibits the previous total e.m.f. E, while the part of the wire between the temperatures T_B' and T_B will now contribute a further e.m.f. given by

$$E' - E = \int_B^{B'} \Delta\sigma dT.$$

Turning to the curves shown in figure 4, we see that the temperature-difference $T_B' - T_B$ as between adjacent curves is not too great to prevent an average tempera-

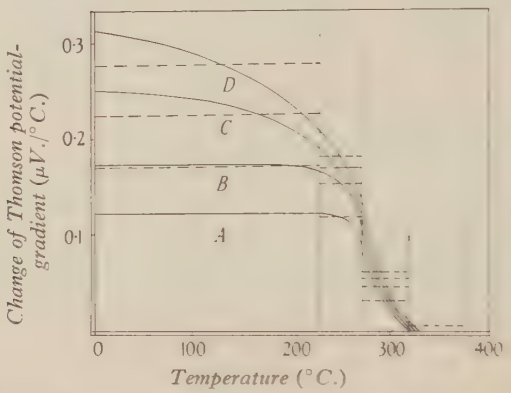


Figure 9. Thomson-potential susceptibility. The average change of Thomson potential-gradien due to fields (*A*) of 1 A., (*B*) of 2 A., (*C*) of 5 A., and (*D*) of 10 A., with the wire under zer tension*. Vertical dotted lines indicate temperature-ranges over which averages were taken, the latter averages being shown by horizontal dotted lines.

$\overline{\Delta\sigma}$ ture from being adopted for the range T_B to T_B', and then an average value $\Delta\sigma$ of $\Delta\sigma$ for that range. We may then consider the change of e.m.f. as we pass from one temperature-distribution to the next, keeping H and S the same, to be given by

$$E' - E = \int_B^{B'} \Delta\sigma dT = \overline{\Delta\sigma}\,(T_B' - T_B).$$

By choosing successive steps we can thus obtain a curve of $\Delta\sigma$ against average temperature $\frac{1}{2}(T_B' + T_B)$ for each value of H and S. The initial step for the whole temperature-range of the lowest temperature is too big to give reliable averages, but it can be treated in the way adopted above for the purpose of providing a rough guide as to the behaviour of $\Delta\sigma$ at lower temperatures. The curves obtained in this way are shown in figure 8 for zero S, and in figure 9 when $S = 8$ kg.

It will be observed that $\Delta\sigma$ for any given H becomes zero near the known Curie point, at least for zero tension. There is an apparent shift of Curie point to higher temperatures when the tension is increased. This, as has already been mentioned, is interpreted as due to the Benedick effect ignored in this first approximation.

* A magnetizing field of 10 amperes corresponds to 303 gauss.

Further analysis in an attempt to get complete and precise dependence of $\Delta\sigma$ on T and of σ on H and S is therefore reserved until data on a and b are available.

The change in Thomson potential-gradient $\Delta\sigma$ (and thus the e.m.f. observed) is taken as positive when the positive current would, were the outside circuit closed, flow up the temperature-gradient in that part of the wire where the magnetic field or tension is causing the change in σ.

§ 5. NOTES ON DEMAGNETIZATION PHENOMENA

In the work on thermomagnetic hysteresis as affected by strain it was noticed that demagnetization destroyed not only the magnetism due to field H, but also that due to tension S. Thus, suppose E_0 to be the initial e.m.f. under zero field, so that $H=0$, and zero tension so that $S=0$. The hysteresis loop against H is obtained for some definite tension, say S_1, and afterwards demagnetization is carried out while $S=S_1$; it is found that E becomes E_1 which is $<E_0$. Releasing the tension, $S \to 0$,

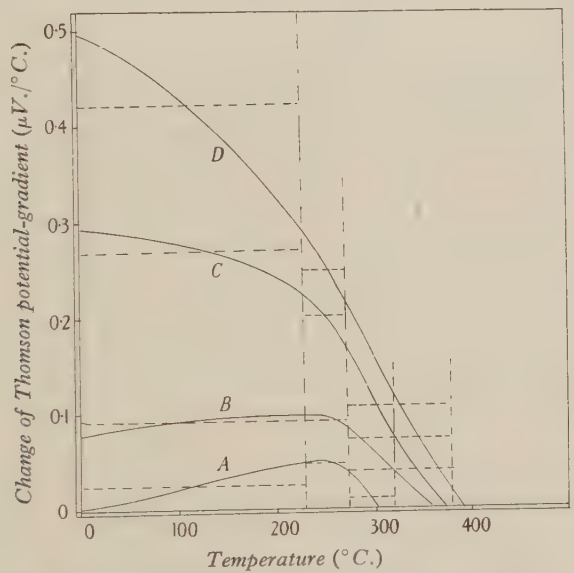


Figure 10. Thomson-potential susceptibility. The constitutions are the same as for figure 9, but the wire is under a tension of 8 kg.*

increases E_2 towards E_0, and a further demagnetization brings E back again to E_0 precisely. For a single demagnetization to give a direct check of E_0 it is necessary to release the tension beforehand.

The magnitudes of E_0-E_1 and E_0-E_2 are of the same order as the values of E shown in figure 3 for the pure thermomechanical e.m.f.

From the above facts we may deduce the following interesting conclusions. (*a*) The radial magnetism produced by S_1 is destroyed by the usual demagnetization even while S_1 is mechanically maintained. (*b*) Even though this magnetism has been destroyed, stretched wire is still thermoelectrically different from unstretched wire;

* A tension of 1 kilogram corresponds to 127·3 kg./cm.^{2}

otherwise E_1 and E_0 should be identical. (*c*) On releasing the tension in the demagnetized wire radial magnetism reappears with sense opposite to that caused ordinarily by that tension: because the change E_1 to E_2 is in the positive direction, whereas the e.m.f. produced by tension is negative.

It was also noticed that the e.m.f. retained after a pure tension hysteresis cycle, figure 3, was removed by the usual demagnetizing process. This might be explained as due to magnetostriction vibrations during demagnetization if the hysteresis is regarded as purely mechanical in origin.

REFERENCES

(1) Heinz Broili. *Ann. d. Physik*, **14**, 3 (1932).
(2) Chang and Band. *Proc. Phys. Soc.* **45**, 4 (1933).
(3) Pi and Band. "Longitudinal thermoelectric effect: (2) nickel in longitudinal magnetizing field." Page 852 of this volume.
(4) Thomson. See the statement of theory in the work on copper. *Proc. Phys. Soc.* **46**, 3 (1934).
(5) D. P. Ray-Chaudhuri. *Zeits. f. Phys.* **71** (7–8), 473–7 (1931).
(6) "Longitudinal thermoelectric effect: (5) silver." Page 904 of this volume.
(7) A. W. Smith. *Rev. Sci. Instr.* **5**, 3 (Oct. 1932).
(8) See for instance curves given for nickel in *Applied Magnetism* by T. F. Wall (E. Benn Ltd.), p. 88.

55.037.2 631.437.1

THE ELECTRICAL PROPERTIES OF SOIL AT FREQUENCIES UP TO 100 MEGACYCLES PER SECOND; WITH A NOTE ON THE RESISTIVITY OF GROUND IN THE UNITED KINGDOM

By R. L. SMITH-ROSE, D.Sc., Ph.D., A.M.I.E.E.,
National Physical Laboratory

Received April 2, 1935. *Read June* 21, 1935

ABSTRACT. Previous papers have described the use of a laboratory method for investigating the electrical properties of soil under alternating-current conditions at frequencies ranging from 50 c. sec. up to 10 Mc. sec. The measurements have now been extended to 100 Mc. sec., and the results are described in the present paper. Substantially the same method has been used, but in order to provide an overlap with the previous work, the present measurements cover the frequency range of 1 to 100 Mc./sec. Check measurements carried out on small fixed resistors indicate that the overall accuracy at the highest frequencies is better than 20 per cent. As a result of this work it may now be stated that for normal samples of surface soil taken from the National Physical Laboratory, the conductivity is of the order of 10^8 e.s.u. at all frequencies up to 1 Mc./sec., rising to rather less than twice this value at 100 Mc. sec. Over the same frequency range, the dielectric constant decreases from about 10^5 at a frequency of 50 c./sec. to about 15 at 100 Mc./sec.

An appendix to the paper draws attention to the electrical-resistivity maps of England and Wales and of Southern Scotland, recently published by the British Electrical and Allied Industries Research Association. These maps indicate in coloured form the apparent resistivity of the ground for a depth of 500 ft. A comparison is made between the values so indicated and those obtained by the author in previously described measurements on samples of soil obtained from various sites in England and Wales.

§ 1. SCOPE OF INVESTIGATION AND METHOD ADOPTED

Considerable attention has been given during the past few years to investigations of the electrical properties of soil under alternating-current conditions. Previous papers[1, 2] have described the use of a laboratory method for measuring these properties and the results obtained on a variety of types of soil over a frequency range from 50 c./sec. to 10 Mc./sec. (wave-length 30 metres). The present paper describes the extension of these measurements to a frequency of 100 Mc./sec. (wave-length 3 metres).

For purposes of clarity it will be an advantage to recapitulate briefly the principles of the method employed. A condenser of suitable size, shape and capacitance is filled with a sample of the soil under examination, and its capacitance and effective shunt resistance are measured at the required frequency. From the values so obtained and a knowledge of the capacitance of the condenser with air as the dielectric, the specific conductivity and dielectric constant of the soil can be ascertained.

§ 2. APPARATUS AND EXPERIMENTAL PROCEDURE

The measurements at frequencies up to 10 Mc. sec. were made by using a concentric cylindrical condenser, a section of which is indicated in figure 1 (*a*). The principal dimensions of this condenser are indicated in the figure, and its capacitance with air as the dielectric was found to be $2{\cdot}5\mu\mu$F. For the higher frequencies the physical dimensions and capacitance of this condenser make its use difficult,

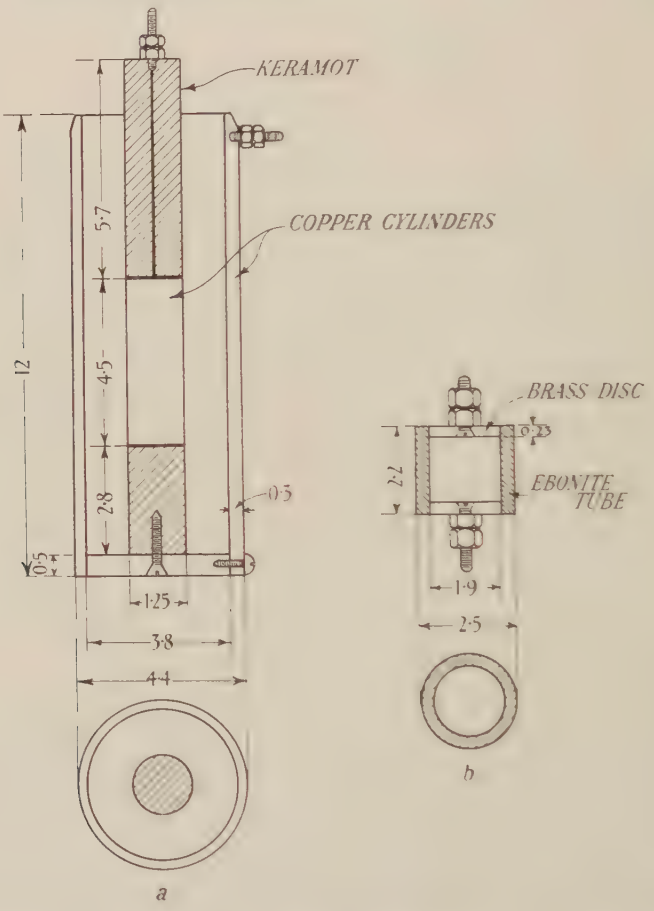


Figure 1. Diagrams of condensers used for soil measurements. (*a*) Large container used for frequencies up to 10 Mc. sec. (*b*) Small container used for frequencies of 10–100 Mc. sec.

and the measurements at frequencies between 10 and 100 Mc. sec. were accordingly made in the small parallel-disc condenser depicted in figure 1 (*b*). The capacitance of this condenser, with air as the dielectric, was calculated to be $0{\cdot}15\,\mu\mu$F.; and this value was checked by measuring the capacitance when the condenser was filled with distilled water.

For measurements on soil the appropriate condenser was filled with the sample under examination, and was connected in parallel with a resonant circuit com

prising a small inductance coil of a few turns and a calibrated variable condenser of maximum capacitance about 20 $\mu\mu$F. At the highest frequency of 100 Mc./sec., the dimensions of the resonant circuit were reduced as much as possible: a single-turn loop 8 cm. in diameter was connected directly across the variable condenser, with all leads reduced to the minimum. The supply of oscillatory current for the measurements was obtained from a valve oscillator of reasonably good frequency-stability. An oscillating detector-amplifier was used to provide, by the heterodyne beat method, a signal audible in a pair of telephones. It was found to be neither readily practicable nor necessary to employ the double-heterodyne-beat method which was used in the measurements at lower frequencies[2].

In carrying out a series of measurements, the tuned circuit alone was first coupled to the oscillator and adjusted to resonance. The variable tuning condenser was then altered on either side of the resonance position to ascertain the capacitance-change required to produce the maximum frequency-change of the oscillator. The effective shunt resistance R_0 of the circuit may then be calculated from the formula R_0

$$R_0 = \frac{1}{\pi f . \Delta C_0} \qquad \text{.....(1),}$$

where ΔC_0 is the total change in capacitance for the maximum frequency-change ΔC_0
and f is the frequency. f

The sample condenser full of soil was then connected across the circuit, and the above operations were repeated to find the new value of capacitance-change, ΔC_1. ΔC_1
The resistance of the condenser full of soil may then be obtained from the relation

$$R_s = 1/\pi f\,(\Delta C_1 - \Delta C_0). \qquad R_s$$

The change in the adjustment of the variable condenser in the two resonance conditions was due partly to the value of the added capacitance and partly, and to a smaller extent, to the effect of the resistance added to the tuned circuit. Allowance could be made for this latter effect so that the capacitance of the condenser full of soil was known, and from this the effective dielectric constant of the sample of soil under examination was obtained.

§ 3. PRECAUTIONS OBSERVED AND CHECKS ON METHOD

It will have been noted from the previous section that the measurements are made in terms of capacitance-changes of a variable air condenser. It is therefore necessary to know with some precision the effective capacitance of the condenser between its terminals. If C is the static or low-frequency capacitance of the con- C
denser, and L is the residual inductance measured from its outside terminals, then L
the effective capacitance C_e at any frequency $\omega/2\pi$ is given by the relation C_e

$$C_e = \frac{C}{1 - \omega^2 LC} \qquad \text{.....(2).}$$

For most ordinary condensers used at low radio frequencies the quantity $\omega^2 LC$ is reasonably small, but at high frequencies even a good-class condenser with a

reasonably low residual inductance may have an effective capacitance which differs considerably from its low-frequency value, particularly when the value of C itself is large. In the case of the larger of the two variable condensers used in these measurements, the value of L was found to be 0·04 μH., so that at maximum scale reading the ratio of C_e/C was 1·13 for a frequency of 60 Mc. sec. This error in the apparent or low-frequency capacitance was allowed for throughout the measurements. The other condenser had such short leads to the inductance coil that their effect was considered to be negligible even at the highest frequency employed.

For the purpose of accurately locating the soil condenser, small mercury cups were provided in the stout leads connecting the variable condenser to its inductance coil. These cups were arranged to receive the very short leads from the terminals of the soil condenser, or alternatively from small fixed resistances used for checking purposes. Care was taken throughout the measurements to avoid the influence of stray capacitances and the effect of any body movements of the observer.

In order to obtain some check of the accuracy of the resistance values obtained by this method at very high frequencies, measurements were made on a number of small fixed resistors of the right order of magnitude. These resistors were of commercial pattern and of four different types, varying from a thin film deposited on a tube 8 mm. in diameter to solid rods of material of from 1 to 10 mm. in diameter. Table 1 shows the results of a series of measurements on these resistors by the method described above. In a few typical cases the values of the resistance measured at 1 Mc./sec. are included.

Table 1. Measurement on four types of resistors at various radio frequencies

Resistor type	Measured values of resistance at frequencies (Mc./sec.) d.c.	1	30	60	100
1	1020	—	930	990	860
	2060	—	1730	1930	1900
2	1000	995	940	970	840
	2070	—	1860	2080	1900
3	1030	1030	950	970	820
	2150	—	1790	2250	2100
4	1030	1030	970	990	910

The largest error in these measurements is 20 per cent, and in the majority of cases the discrepancy between the d.c. and a.c. values is less than 10 per cent. This discrepancy includes both the error of the measurement and any change in value of the resistor at high radio frequencies from its direct current value. Since, as has already been remarked, the resistors were of very different types, they would not be expected to show a consistent change in their values at the higher frequencies, which might influence the above results in the same direction for each type. As it was considered that an accuracy of 20 per cent was adequate for the measurements on soil, the above discrepancies were not investigated in any great detail.

It is to be observed in the above table that there is a tendency for the radio frequency value of the resistance to be less than the direct current value. In the

case of a resistor of type 4 having a direct-current resistance of 1990 Ω., the value obtained at the high radio frequencies was about half this figure. This discrepancy was traced to the fact that a resistance film was deposited on both the inner and outer walls of a porcelain tube, but that conductive connexion from the terminals was made only to the outer film. At very high frequencies, however, the effect of the capacitance between the films was to connect them in parallel with each other, so as to give an effective resistance approaching half the value obtained in a direct-current or low-frequency measurement. On removal of the inner film, the value of the resistance was found to be 1930 Ω. at the frequency of 60 Mc./sec. It is possible that some of the other errors recorded in the above table are due to parallel-resistance effects of an analogous type.

§ 4. MEASUREMENTS ON SOIL

By means of the experimental procedure outlined above, measurements were made on samples of soil taken from a site at the National Physical Laboratory, which had been used as site 1 in previous work[1]. Some check measurements were carried out at a frequency of 10^6 c./sec. in order to establish continuity with the results recorded in previous papers. In these cases the measurements were made both in the small condenser used for the present frequency range and also in the larger condenser used in all the previous work. Some care was necessary in packing the smaller condenser to the required degree, but the agreement obtained between measurements at the same frequency in the two condensers indicates that this packing was satisfactorily accomplished. As in the previous work, the moisture-content of the soil was determined from the loss of weight on drying it out at a temperature of 100° C. The values of the moisture-content included in table 2 are referred to the weight of the dry soil.

Table 2. Summary of measurements on soil from site no. 1 at the National Physical Laboratory

Sample No.	Moisture-content (per cent)	Frequency (Mc./sec.)	Conductivity*		Dielectric constant	
			In small container (e.s.u. × 10^{-8})	In large container (e.s.u. × 10^{-8})	In small container	In large container
A	24	1·0	0·9	1·0	55	55
		10	1·3	—	25	—
		60	1·3	—	18	—
		100	1·5	—	15	—
B	25	1·0	0·9	0·9	40	50
		10	1·1	—	21	—
		30	1·3	—	21	—
		60	1·5	—	20	—
		100	1·9	—	15	—
C	30	1·0	0·8	0·8	45	50
		10	1·0	—	25	—
		30	1·2	—	20	—
		100	1·3	—	20	—

* Conductivity (e.s.u.) = $9 \times 10^{11} \div$ resistivity (ohm-cm.).

These results show that for the samples of soil selected the conductivity increases progressively from about $0{\cdot}9 \times 10^8$ e.s.u. at a frequency of 1 Mc. sec. to a value of from $1{\cdot}3$ to $1{\cdot}9 \times 10^8$ e.s.u. at 100 Mc. sec. Over the same frequency range the dielectric constant falls from about 50 to a limiting value which appears to lie between 15 and 20. The agreement between the results obtained with the two condensers at a frequency of 1 Mc. sec. is considered to be quite satisfactory and to justify the linking up of the present results with those previously obtained at low frequencies. These showed that the value of the conductivity remained practically unchanged for all frequencies from 50 c. sec. up to 1 Mc. sec., although the dielectric constant decreases steadily from very high values at low audio frequencies.

§ 5. CONCLUSIONS

It may thus be concluded that for soil of normal moisture-content from the particular site selected at the National Physical Laboratory, the conductivity is of the order of 10^8 e.s.u. at all frequencies up to 1 Mc. sec., and then rises to rather less than twice this value at 100 Mc. sec. Over the same frequency range the dielectric constant decreases from about 10^5 at a frequency of 50 c. sec. to about 15 at 100 Mc./sec. The abnormally high values experienced at the lower frequencies have been discussed briefly in a previous paper[2]; they are probably attributable to the formation of a polarization film over the surface of the electrodes. In the case of soil, these high values of dielectric constant are of no practical importance at the frequencies at which they are obtained. It is only at frequencies above 1 Mc./sec. that the dielectric constant need be taken into account in most radio problems.

The variation of these electrical properties with the nature of the soil as obtained from different localities, with moisture-content and temperature has been previously investigated in some detail at frequencies up to 10 Mc. sec.[2]. It has not been considered useful at this stage to extend this study to the higher frequencies now covered.

§ 6. ACKNOWLEDGMENTS

This investigation was carried out as part of the programme of the Radio Research Board, and the paper is published by permission of the Department of Scientific and Industrial Research. The author is indebted to the able assistance rendered by Mr A. C. Gordon-Smith in connexion with measurements involved in the work.

REFERENCES

(1) R. L. Smith-Rose. "The electrical properties of soil for alternating currents at radio frequencies", *Proc. roy. Soc.* A, **140**, 359–77 (1933).

(2) R. L. Smith-Rose. "Electrical measurements on soil with alternating currents", *J. Instn. elect. Engrs.* **75**, 221–37 (1934).

APPENDIX

NOTE ON THE ELECTRICAL RESISTIVITY OF GROUND IN THE UNITED KINGDOM

The British Electrical and Allied Industries Research Association have recently published* electrical-resistivity maps of England and Wales and of Southern Scotland. These maps indicate in coloured form the apparent resistivity of the ground for a depth of 500 ft., and they have been prepared by Mr A. Broughton Edge for the above Association. During the past year or two the author has been investigating by a laboratory method the electrical properties of soil for alternating-current conditions at various frequencies from 50 c./sec. up to high radio frequencies. The results of such measurements made on samples of soil taken at various depths down to 10 ft. (in one case 300 ft.) from some 16 different sites in England and Wales were published in 1934†. It is interesting to compare these results with those indicated on the electrical-resistivity map referred to above.

The following table shows such a comparison for the 16 localities in question. The author has utilized the results obtained by him at a frequency of 1000 c./sec.,

Electrical resistivity (kΩ.-cm.) of soil in different parts of England and Scotland

Site	Value according to the map by A. Broughton Edge for depth of 500 ft.	Values obtained by R. L. Smith-Rose for depth of 10 ft.
Rugby, Warwick	1 to 3	1·5
Baldock, Herts	3 to 30	6·9
Tatsfield, Kent	3 to 30	6·0
Brookmans Park, Herts	3 to 30	6·0
Daventry, Northants	1 to 3	35
Washford Cross, Somerset	1 to 3	2·1
Brendon Hills, Somerset	30 to 300	$3{\cdot}5 \times 10^4$
Merrivale, Devon	>300	6×10^5
Dousland, Devon	>300	3×10^5
Moorside Edge, Yorks	3 to 30	29
Westerglen, Falkirk	3 to 30	4·3
Teddington, Middlesex	<1	7·5
Droitwich, Worcester	1 to 3	2·4
Slough, Bucks	1 to 3	4·0‡
Banstead, Surrey	3 to 30	4·5‡
Pulham, Norfolk	3 to 30	3·5‡

‡ Soil taken within 1 foot of surface and measured at 100 kc./sec.

and it is understood that the measurements by Mr Broughton Edge were made under direct-current conditions or with alternating-currents of various frequencies up to about 3000 c./sec. Measurements published in the paper referred to above have shown that, except in the case of the higher-resistivity samples of stone and

* Shown at the Physical Society's Exhibition, January 1935.

† See reference (2) above.

slate, there is no substantial change in resistivity of soil in changing the frequency from 50 to 1000 c./sec.

The table shows that in 12 cases out of the 16 cited there is good agreement between the two sets of measurements. In these 12 cases, moreover, the author's results show that except for surface soil, the values of resistivity obtained at various depths from 1 to 10 ft., all lie within the limits specified by the map. The note on the map stating that "it is unsuitable for obtaining the resistivity corresponding to shallow depths" may thus need some qualification.

The cases in which there is a discrepancy between the two sets of results are discussed individually below.

Daventry. There is a 10 : 1 discrepancy in the values of resistivity indicated in the table. It may be remarked that the ordinary geological map* shows streaks of middle and upper lias through the Daventry and Northampton region, as distinct from the lower lias at Rugby. This is reflected in the author's measurements, which show that the resistivity at Daventry, where the subsoil is sandy, is over 20 times as great as that at Rugby, where the subsoil is clay. No similar streaks appear on the B.E.A.I.R.A. map, which shows a large belt of the midlands in the same resistivity range.

Brendon Hills. In this case the author's measurements on all samples taken from depths of 1 to 10 ft. showed a resistivity greatly in excess of 300 kΩ.-cm., the maximum value ascribed to this region by the B.E.A.I.R.A. map. The samples comprised a mixture of loam and slate, the slate-content and resistivity increasing with the depth. It may be noted that for this part of the country the resistivity map follows very closely the geological map to which reference was made above.

Teddington. Measurements of the resistivity of the ground within a few feet of the surface at the National Physical Laboratory, Teddington, have been carried out for several years past by a variety of methods and at various frequencies from 50 c./sec. up to 10^8 c./sec. In all cases the results have indicated a resistivity in excess of 2 kΩ.-cm., and generally above 5 kΩ.-cm. The probable explanation of the much lower value obtained by Broughton Edge is that it is due to the influence of the relatively low-resistivity clay bed which would be encountered at greater depths.

Slough. A similar effect may be influencing the results at Slough, although to a lesser extent on account of the proximity of this locality to the edge of the clay bed; for here again measurements of the ground resistivity for small depths at the Radio Research Station, Slough, have consistently given values in excess of 3 kΩ.-cm.

Conclusion. From this survey it appears quite definitely that the electrical resistivity map compiled by A. Broughton Edge and recently published by the B.E.A.I.R.A. is a valuable contribution to our knowledge of the resistivity of the ground in this country. It is possible that a reference to the detailed observations from which the map was compiled might elucidate one or two of the discrepancies pointed out above. Further, the available evidence shows that the map forms a safe guide to the resistivity values for alternating-currents of all frequencies from

* Geological Map of England and Wales by Sir Archibald Geikie.

50 c./sec. into audio range. Although the map is intended to indicate the resistivity corresponding to depths of about 500 ft., it would appear in many localities to be fairly representative even for depths of only a few feet. Further work on an extended scale would, however, be necessary before a definite recommendation could be made that the map should be used as a reliable guide to surface conditions.

DISCUSSION

Mr G. D. Pegler. The author's observation on resistors should be of value to workers at high frequencies. Since a comparison is made between the values of soil-resistivity measured by him for depths down to 10 ft., and the results indicated on the map for a depth of 500 ft., it would be interesting to know if he has made any investigation of the variation of soil-resistivity with pressure.

Dr L. Hartshorn. I am inclined to think that the low values of resistance obtained for the fixed resistors at very high frequencies are not altogether due to experimental error. I have observed such an effect even in resistors of the solid-rod type, and it is to be expected in all composite materials since they may be regarded as consisting of a complicated network of resistances and capacitances. The conductance of any such combination must increase with increasing frequency. The effect is well known in solid dielectrics. The soil and the composite resistor may be regarded as special cases, in which the effect becomes considerable in the range of frequencies used in this investigation.

Author's reply: In reply to Mr Pegler: I have in the course of my investigations on the electrical properties of soil made certain measurements on the effect of pressure on these properties. These measurements were, however, carried out merely with the object of ensuring that the laboratory tests should be representative of the actual conditions of the soil at the site under exploration. In the case of all results published, the sample condensers have been packed to a density approximately equal to that of the soil in its natural state.

Dr Hartshorn's comments are gratifying in suggesting that the experimental error is possibly somewhat less than that indicated in the paper.

548·73

THE QUANTITATIVE MEASUREMENT OF THE INTENSITY OF X-RAY REFLECTIONS FROM CRYSTALLINE POWDERS

By J. C. M. BRENTANO, D.Sc., F.Inst.P.

Received January 29, 1935. *Read in title May* 3, 1935.

ABSTRACT. The paper refers to determinations of the dispersion of F-values in the range of the L absorption levels, in which powder methods were applied to measurements requiring a high degree of accuracy and discusses the conditions determining quantitative measurements made with microcrystalline powders.

The absorption in a powder layer and the effect of the size of the crystallites on the evaluation of intensities are discussed. Indications are made as to the preparation of suitable particles and as to the use of spacing materials for avoiding the formation of larger units. A detailed discussion is given of the factors determining quantitative measurements by the method of the flat powder layer and by the method of mixed powders. The conditions determining the accuracy of the microphotometric photometry of reflections obtained from a flat layer and the absorption effect in a mixed layer in relation to the absorption in the individual particle are considered. Expressions are derived for evaluating the absorption effect for particles of different absorbing-power.

Reference is made to a recent paper by Brindley and Spiers on a development of the method of the flat powder layer and of the method of mixed layers. It is shown that this particular procedure which had been described by the writer in an earlier paper is not advisable for accurate measurements.

§ 1. INTRODUCTION

Recently a series of measurements have been made on the dispersion of X-ray scattering in the region of the L absorption levels by Baxter and the author[1], leading to the establishment of an F-dispersion curve for tungsten for the wave-length range from 0·49 A. to 2·3 A. The evaluation of the dispersion which was derived from the differences of F-values obtained for different wave-lengths, required a high degree of accuracy; on the other hand particular difficulties were presented by the absorption in the tungsten, which is considerable for some of the wave-lengths used, and by the fact that the measurements had to be made for strong reflections in the range of small values of $\sin\theta/\lambda$ where extinction effects are particularly large.

The measurements were made upon powders by determining photographically the intensities of the reflections scattered from a flat layer and using a mixture of powders[2], so as to obtain relative measurements in terms of a reference substance. These two methods, introduced by the writer in 1928 for quantitative evaluation from powders, have since been used in a number of instances. Intensity-measure-

ments from powders present considerable difficulties; on the other hand they are the only way of approach to a number of determinations where large single crystals cannot be used, and we believe that a remarkably high degree of accuracy can be obtained if due precautions are taken. Discrepancies in particular with the method of powder mixtures in the hands of different observers seem to be due to the fact that proper regard had not been given to the conditions requisite to such measurements. A more comprehensive analysis should therefore be given, as the conditions determining the accuracy of powder measurements have only been discussed in an incidental way.

Brindley and Spiers have recently discussed in these *Proceedings*[3] a particular development of the methods. This had been described by the author in an earlier paper that probably escaped their attention; reference to this arrangement, which is not advisable for accurate measurements, will be made in the course of this paper.

§ 2. EXTINCTION EFFECTS IN POWDERS

One particular difference between the use of fine crystalline powders and single crystals for measuring the intensities of X-ray reflections depends on the fact that the size of the powder particles determines an upper limit to the size of the regularly spaced crystal units or crystallites. In this way by using a sufficiently fine powder the extinction effect of large single crystals, which Darwin describes as primary extinction and which in the first place impairs the interpretation of the intensities of reflections from large crystals, can be avoided. These conditions have been discussed by Darwin[4] and Ewald[5] and need not be recalled here.

Further, the coefficient of absorption enters into the expression for the scattered intensity. This has a different character for a fine powder and for an extended single crystal of mosaic structure. For the latter the absorption coefficient for a given wave-length in an angular range near the angle of reflection contains a term for secondary extinction, which increases as the setting for a strong reflection is approached in a manner depending on the angular spread of the blocks constituting the particular mosaic. For a powder of sufficiently small crystallites oriented at random an effect analogous to secondary extinction takes place in so far as the beam of wave-length λ going through the powder encounters crystallites so oriented as to satisfy the reflecting condition λ

$$\sin\theta = \lambda/2d_{(hkl)} \qquad \text{.....(1)},$$

where θ is the glancing angle of incidence and of reflection with respect to a set of lattice planes (hkl) with the spacing $d_{(hkl)}$. All lattice spacings for which the spacing $d_{(hkl)} > \lambda/2$ contribute to this effect, which we have shown[2] to be a discontinuous function of the wave-length; with powder particles oriented at random it is however the same for all directions and independent of the particular reflection we observe. The powder offers thus the advantage that the same absorption coefficient enters into the intensity expression for all reflections and further that a greater number of reflections can be observed than with a single crystal, where the measurement is dependent on the presence of sufficiently well-developed crystal faces. θ $d_{(hkl)}$

On the other hand the numerical value of the absorption coefficient varies greatly with the individual powder specimen; it depends on the closeness of the packing and on the extent to which the interspaces between the powder particles are filled with binding or spacing material; further the appearance of the full number of reflections leads to superpositions.

These points largely determine the procedure to be adopted in the interpretation of intensity-measurements from powders.

§ 3. SCATTERING FROM VOLUME ELEMENT OF POWDER

In discussing the X-ray scattering from a micro-crystalline powder it is convenient to refer to the intensity of reflection from a volume element dV sufficiently small to produce no appreciable absorption and containing an adequate number of crystallites oriented at random so as to present no preferential orientation of lattice planes with respect to any direction.

The intensity of scattering is then given by

$$dP_{(hkl)} = \tfrac{1}{2} I p_{(hkl)} F^2_{(hkl)} N^2 q \cos\theta dV \qquad \text{.....(2),}$$

where $P_{(hkl)}$ is the total energy scattered in unit time for the reflection (hkl) out of a parallel incident beam for which I represents the intensity, i.e. the energy per unit time through unit cross section. $F_{(hkl)}$ is the scattering factor per unit cell (often the letter S is used for this term to distinguish it from the scattering factor of the atom) and N is the number of scattering cells per unit volume of the powder. θ is the glancing angle relating to the particular reflection and wave-length. $p_{(hkl)}$ is the multiplicity factor of the reflection, that is, the number of times the particular lattice appears with different orientation in the crystal form. The quantity q, where

$$q = \frac{e^4}{m^2 c^4} \frac{\lambda^3}{\sin 2\theta} \cdot \frac{1+\cos^2 2\theta}{2} e^{-b\frac{\sin^2\theta}{\lambda^2}},$$

collects those quantities which establish the numerical link between I and $dP_{(hkl)}$ on the basis of classical theory, in accordance with the usual definition of the scattering-factor F. The exponential term $e^{-b\frac{\sin^2\theta}{\lambda^2}}$ relates to heat motion. No indices are given to θ and q, but they also refer to the particular reflection.

The scattering-factor $F_{(hkl)}$ which defines in the traditional way the scattering from the cell in terms of the scattering from a single free electron depends on the atomic scattering-properties of the atoms constituting the cell and on their spacial arrangement within the cell, leading to additive and subtractive interference effects. It takes account of the fact that only the coherent part of the scattered radiation contributes to the reflection and of the dispersion of X-ray scattering, by which the scattering from any atom depends on the relation between the wave-length and its energy levels. $F_{(hkl)}$ considered as an experimental quantity has no intrinsic value unless it is related to a certain wave-length.

In the case of crystals constituted of one kind of atoms only, as, for instance, metals, F can be introduced as the scattering-factor of the atom, when N becomes

Loschmidt's constant, the number of atoms in unit volume. Only those reflections of the simple lattice will appear for which the scattering contributions from the different atoms are additive according to the type of unit cell of the crystal. In our paper in 1928 the expression was applied to reflections from gold, silver and aluminium and was used in this sense.

§ 4. ABSORPTION EFFECTS IN FINE POWDERS

The scattering from an extended block of powder is obtained by integrating over its volume after multiplying the contributions from each volume-element by an absorption term

$$\exp(-\mu s_{xyz}) \qquad \text{.....(3),}$$

where $s_{(xyz)}$ is the total distance which the incident and scattered beam traverse in the powder as determined by the position of the point xyz within the powder. μ is the linear absorption coefficient of the powder; it acquires a definite value only when the specimen is constituted of such small particles distributed at random that it can be regarded as homogeneous for dimensions in which absorption becomes appreciable. We can then introduce a general coefficient of absorption μ which is related to the absorption coefficient of the constituents according to the volume they occupy.

$s_{(xyz)}$ μ

It can easily be shown that for two constituents 1 and 2 with the coefficients of absorption μ_1 and μ_2 we have

$$\mu = \frac{\mu_1 m_1 \rho_2/m_2 \rho_1 + \mu_2}{m_1 \rho_2/m_2 \rho_1 + 1} \qquad \text{.....(4),}$$

where m_1/m_2 is the ratio of masses and ρ_1/ρ_2 the ratio of densities.

The value of μ differs considerably from the absorption coefficient of the individual crystal measured within or outside the reflecting position and varies considerably with the constitution of the layer; for a given wave-length it will, however, have the same value for all reflections. The coefficient of absorption μ comprises any secondary extinction effect of the crystalline powder particles, to which we have referred.

When the intensities of scattering for various angles for one kind of crystals are compared, for instance when an F curve expressing the change of F with $\sin\theta/\lambda$ for a given wave-length is being determined, the uncertainty about the numerical value of μ for the powder can be overcome by choosing such conditions that for all angles of scattering the absorption term enters into the intensity expression as a constant multiplicative factor. This condition is not generally satisfied with the thin powder rod used in the Debye-Scherrer camera, where the scattering in different directions depends on the extent to which the deeper layers of the rod contribute to the scattered intensity. Definite conditions are obtained with the powder rod only when the absorption in the rod either is negligible or is so great that scattering is confined to a thin layer near to the surface. In the latter case for a given shape of rod the geometric location of the scattering volume becomes definite,

so that the intensity-distribution can be calculated or established experimentally in a general way. This procedure, which is particularly adapted to very heavy atoms, has been introduced by Claassen[6], and has been discussed in detail by Rusterholz[7]. When the depth of penetration into the layer has to be considered, approximate expressions apply, which depend, however, on the numerical value of the absorption coefficient.

§ 5. THE FLAT POWDER LAYER

The writer has discussed the reflection from a flat powder-layer[8,2] and has shown that it can be used in a general way to obtain measurements of relative intensities independent of the absorption coefficient μ.

For a flat layer sufficiently thick to absorb fully the incident radiation, placed so that its normal is in one plane with the direction of incidence and with the direction in which the scattered beam is observed,

$$dP_{\zeta(hkl)} = \frac{Ip_{(hkl)}\, F^2_{(hkl)}\, N^2\, q\, f\, d\zeta}{8\pi\mu\,(1+\sin\alpha/\sin\beta)} \qquad \text{.....(5).}$$

$P_{\zeta(hkl)}$ where $P_{\zeta(hkl)}$ is the quantity of radiation scattered for the reflection (hkl) in unit
ζ time into the small azimuthal angle $d\zeta$ near the plane in which scattering is observed,
f, α, β when the parallel beam of intensity I and cross-section f falls on the layer. α, β are the glancing angles of the incident and scattered beam with respect to the surface of the powder, so that

$$\beta = 2\theta - \alpha \qquad \text{.....(6).}$$

a To meet the actual experimental conditions we consider a beam diverging from the anticathode or from a slit A, figure 1, situated at a distance a from the centre of the layer. The scattered radiation is recorded on a photographic film in

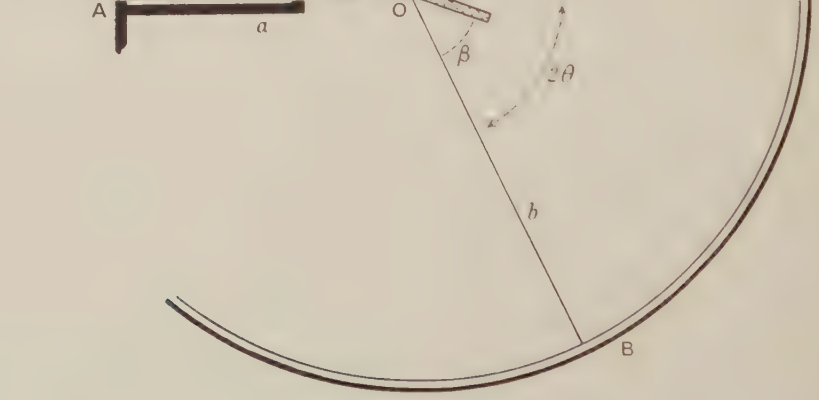


Figure 1.

3, placed on a circle with radius b, having O as centre, so that the X-rays fall normally on the photographic film. For this arrangement

$$P_{b\,(hkl)} = X \frac{p_{(hkl)} F^2_{(hkl)} N^2 q \chi \sin \beta}{16\pi b \sin^2 \theta \mu \cos \frac{1}{2}(\beta - \alpha)} \qquad \text{.....(7),}$$

s the total intensity of the line representing the reflection (hkl) on the film, when a beam subtending the solid angle χ of intensity X per unit solid angle is incident on the layer.* The absorption coefficient enters into equation (7) in $1/\mu$ and is therefore a constant for all reflections from the particular layer. The quantities X, χ and b have common values for the exposure or are constants of the instrument and N is a constant for the particular type of crystal. The link between relative F values and observed P_b values is thus given by

$$P_{b\,(hkl)} = \text{const.} \frac{p_{(hkl)} F^2_{(hkl)} q_{(hkl)}}{\sin^2 \theta} \frac{\sin \beta}{\cos \frac{1}{2}(\beta - \alpha)} \qquad \text{.....(7').}$$

The values of p, q and θ are particular to each reflection and depend on crystallographic data of the lattice only, while the effect of the orientation of the layer with regard to the angles of incidence α and of scattering β is determined by the factor i where

$$i = \frac{\sin \beta}{\cos \frac{1}{2}(\beta - \alpha)} \qquad \text{.....(8),}$$

which by introducing equation (6) can be written

$$i = \frac{\sin (2\theta - \alpha)}{\cos (\theta - \alpha)} \qquad \text{.....(8').}$$

It will be seen that with other conditions equal the intensity for a given reflection, i.e. for a given angle θ, increases with β and with the difference $(\beta - \alpha)$. For small angles the intensities are approximately proportional to β.

§ 6. THE TECHNIQUE OF INTENSITY MEASUREMENTS WITH A FLAT LAYER

Expressions (7'), (8) and (8') can be applied in different ways to the evaluation of intensities.

One way is to make use of a focusing condition of the flat layer, which is established for $\sin \alpha / \sin \beta = a/b$, when sharp lines are obtained on the film in B. A method based on this property in which the powder layer is rotated and a rotating screen is used in front of the photographic film, so as to satisfy the focusing condition over an extended angular range, has been described by the author in earlier papers[2,8]. For this case we can write equation (7') in the form

$$P_{b\,(hkl)} = X \frac{p_{(hkl)} F^2_{(hkl)} N^2 q_{(hkl)} \chi}{8\pi b \sin \theta} \frac{1}{\mu} \frac{1}{1 + \sin \alpha / \sin \beta} \qquad \text{.....(9).}$$

* Expression (7) can be derived from

$$P_b = X \frac{p F^2 N^2 q \chi}{8\pi b \sin \theta} \frac{1}{\mu} \frac{1}{1 + \sin \alpha / \sin \beta}$$

introduced in an earlier paper[2] by considering that

$$\frac{1}{1 + \sin \alpha / \sin \beta} = \frac{\sin \beta}{2 \sin \theta \cos \frac{1}{2}(\beta - \alpha)}.$$

The term $1/(1+\sin\alpha/\sin\beta)$ in (9) takes the place of the term i in the previous expression and is constant so long as the focusing condition is satisfied for a given ratio a/b. This term is thus eliminated in relative measurements.

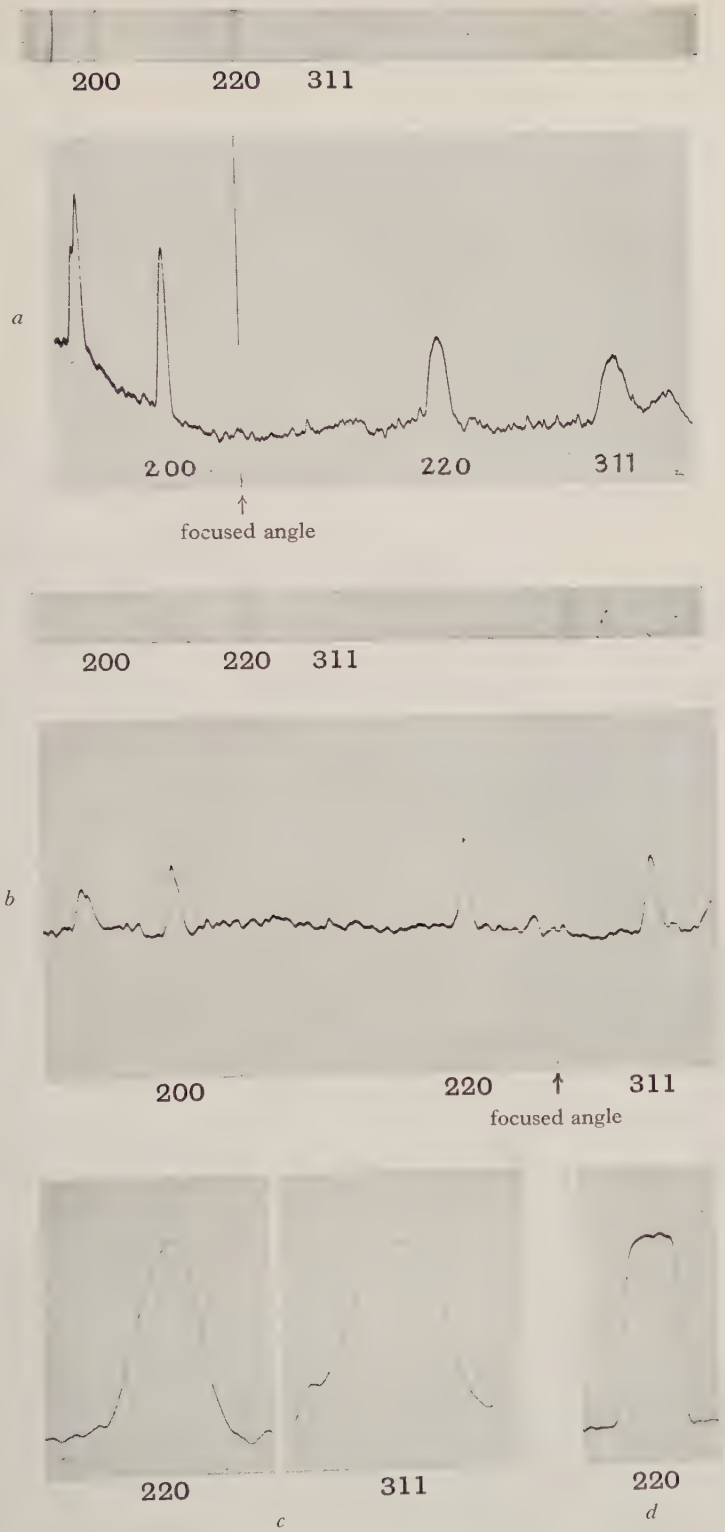
To obtain the gain of intensity derived from an increase of angle β over α as given by equation (8), and at the same time satisfy the focusing condition, the distance a should be made smaller than $b^{(9)}$.*

The other way is to set the powder at a certain angle α with respect to the incident beam and to associate with each value of P_b the appropriate i term introducing in equation (8′) the angle θ belonging to each individual reflection. This method is particularly suitable for comparing lines situated within a small angular range. The angle α is then so chosen as to satisfy the focusing condition for an intermediate point of the range. This procedure has been adopted in the measurement on the dispersion of the scattering-power of tungsten. By choosing the focused angle to be nearer to one or the other line, the sharpness and density of one line can be increased as compared with another. This effect can be used within limits to adjust the densities of unequal lines.

The arrangement making use of a fixed angle α can also be used to compare lines comprising a wider angular range of θ outside the focused angle, and we have discussed this as applied to measurements on aluminium in a paper published in 1931[10]. Measurements were then made comprising the same reflections, with the powder set at the fixed angles $\alpha=25{\cdot}0^\circ$ and $\alpha=35{\cdot}5^\circ$ in order to make possible an internal check by comparison of corresponding reflections in each case. Reference may be made to these exposures in order to indicate some points which arise in this connexion. Figure 2*a* is a reproduction of a film and of its microphotometric record† with the powder set so that $\alpha=25{\cdot}0^\circ$, figure 2*b* relates to an exposure with the powder set so that $\alpha=35{\cdot}5^\circ$. With $a=b$ the focused angle was 50° for the first and 71° for the second. Different reflections were therefore near the focused angle in each case and this accounts for the different broadening of the lines. For the purpose of examining more closely the sharper lines in both exposures, the relevant data are collected in table 1. Column 5 gives the relative i-values of corresponding lines for the two settings, to which the intensities of the reflections should be proportionate. When these are compared with the observed P_b values in column 6, expressed in similar relative units, the agreement is reasonably good, having regard to the quality of some of the lines. It will however readily be seen that the lines are not suitable for exact comparisons owing to the varying spread and intensity dis-

* It should be noted however that the focusing condition introduces a geometric-optical relation between the entrance slit in A and the line in B, so that for a given reflection a linear magnification between slit and line exists in the ratio $a:b$. By this relation the specific intensity of the reflections on to the film is reduced when b increases over a. This does not, however, apply when the actual width of the line is determined by factors other than the magnification of the slit, such as the penetration of the rays into the layer or the imperfection of the crystallites. Whether it is of advantage to make b larger than a depends thus on the particular conditions. In some of the measurements on tungsten a large distance b of the film from the powder was required in order to obtain sufficient separation of the lines, and the width of the lines was essentially determined by the broadening resulting from the smallness of the crystallites. Here it was found of advantage to make a smaller than b.

† The dense films chosen for better reproduction are not those actually evaluated by scattering[14].



ure 2. Scattering of Cu K_α radiation from stationary aluminium powder layer set at $\alpha = 25{\cdot}0°$, figure 2 *a* and at $35{\cdot}5°$, figure 2 *b*. ↑ denotes the focused angle. Comparative records of reflections taken outside the focused angle figure 2 *c*; in focus with wide slit figure 2 *d*.

tribution, according to their distance from the focused angle. This is shown more distinctly by the photometric records, in which the ratios of the peak values differ considerably from the ratios of i-values in column 5, indicating that focusing plays a predominant part in determining the peak values and the density-distribution of the lines.

Table 1. Values of i or $\sin(2\theta-\alpha)/\cos(\theta-\alpha)$ for the settings $\alpha=25{\cdot}0^\circ$ and $\alpha=35{\cdot}5^\circ$ and various reflections from aluminium. Ratios of i and of P_b for the two settings (relative values).

Index	θ	i_{25°	$i_{35{\cdot}5^\circ}$	$\frac{i_{25^\circ}}{i_{35{\cdot}5^\circ}}$	$\frac{P_{b25^\circ}}{P_{b35{\cdot}5^\circ}}$
200	22·2°	0·333	0·159	2·10	2·10
220	32·4°	0·646	0·490	1·32	1·39
311	38·8°	0·818	0·673	1·21	1·10

§ 7. THE EFFECT OF FOCUSING

The distribution of the density across the line has a considerable bearing on the accuracy of the measurement when photographic methods are employed. The accuracy decreases when large and small densities have to be photometered side by side and it is desirable that the lines to be compared should have similar densities. Errors do not arise so much in the evaluation of levels of different density, represented by steps of adequate area when a blackening curve can be introduced with a high degree of certainty, as in the evaluation of an area of varying density when the density of each element has to be associated point to point with the appropriate intensity-value. This is particularly felt when relevant parts of the line extend through a range of widely varying densities. It is therefore desirable to obtain lines with a central part of approximately uniform density and a steep falling off on both sides to the density of the background. With crystallites giving sufficient definition an approach to this type of line can be obtained with a flat powder layer by satisfying approximately the focusing condition and using a wide entrance slit in A; the line will then be an image of this slit. Figure 2*d* is the microphotometric record of the aluminium (220) reflection obtained in this way, with an entrance slit 0·8 mm. wide. Figure 2*c* is a microphotometric record at equal magnification of the (220) and (311) reflections as in figure 2*a*. It will be seen that in this case the gradual decrease of density is not suitable for accurate measurement. It is for this reason that when we are recording a group of lines simultaneously with one fixed setting α of the layer we prefer to embrace only a limited angular range of about 20° for small angles and of 50° for large angles of deflection, and to use the rotating screen for larger angular ranges, when high accuracy is required.

Brindley and Spiers[3] recently described and discussed the use of a flat layer for quantitative intensity-measurements without having apparently noticed our earlier work on the subject. They deduce an intensity expression similar to our

term (8′) and apply this to a layer set at a fixed angle α, observing the lines for different reflections in the same way as we have done in the measurements on aluminium to which we referred. They also discuss again the method of powder mixtures, which we described and used throughout in our earlier work.

It is very welcome to see the method of flat powders discussed from various points of view and the paper of Brindley and Spiers has the special merit of giving prominence to the effect of the angular term, which they call the absorption term, in determining the relative intensities of reflections at various angles. They indicate in a graph the change of it with angle α and refer to this as a means for controlling the relative densities of lines situated at different angles. It should be noted however for practical application that the i term affects the total intensity of the line as represented by P_b, while the varying broadening of the lines according to their position with regard to the focused angle (which is a distinctly different effect) is the major factor in determining the peak values and the density-distribution of the lines. We differ from Brindley and Spiers in considering that this broadening is not favourable for good measurements when it leads to considerably enlarged lines. The broadening outside the focused angle is reduced and i becomes the deciding factor when very narrow beams are used. This however reduces not only the intensities, but also the area of powder used for scattering, which impairs the averaging and is a disadvantage, particularly when the powder is held stationary as suggested by Brindley and Spiers.

With very small particles or with particles constituted of light atoms, when the X-rays penetrate to a greater depth and a greater number contribute to scattering, a stationary layer can be used; and such a layer can also be used with comparative immunity for large angles of deflection θ, when the penetration in the layer is again increased and a greater number of particles contribute to scattering. We have shown[8] that the broadening of a line outside the focused angle is proportional to $\epsilon\eta \cot \beta\, a/(a+b)$, where ϵ is the angular width of the incident beam in the plane AOB and $\eta/2$ the angular distance of the reflection from the focused angle. This broadening effect due to departure from the focusing condition is thus very much reduced for large angles and lines situated over a much wider range can be compared under conditions of approximate focusing. Measurements at large angles are generally undertaken under much more favourable conditions. Their use is limited because of superposition effects due to the crowding of many reflections when mixtures of different substances are used, so that it is difficult to determine the correct background. In many cases also the evaluation of F-values for small angles is of particular interest.

§ 8. THE EFFECT OF THE SIZE OF CRYSTALLITES

Another factor which affects the density-distribution depends on the limited resolving power of small or distorted crystallites comprising only a limited number of regularly spaced lattice planes. Broadening of the lines becomes noticeable when the number of regularly spaced planes is less than 1000 and increases with decreasing

number. The intensity-distribution of the lines depends then on the size-distribution of the scattering elements. Side by side with the condition that crystallites must not be too large, imposed by extinction, we have thus the requirement that the crystallites should not be too small. Particularly unfavourable conditions are found for a powder containing a mixture of larger and smaller crystallites extending into the range of very small units, which we described as a vanishing size-distribution; this has been discussed in another paper[10]. But apart from this case a powder consisting of small crystallites gives an unfavourable falling off of intensity which is undesirable for accurate evaluations.

This is not the place to discuss the preparation of suitable powders in detail; it may be recalled from previous papers[1, 2, 10], however, that the process of mechanical grinding is generally inadequate: the particles obtained are too coarse, at the same time they embody, particularly in the case of ductile metals, a certain number of highly deformed crystallites presenting all the characteristics of a vanishing size-distribution. Chemical precipitation and controlled processes of chemical reduction have proved successful in a number of instances, and here it was often found useful to introduce a spacing material to avoid the formation of larger units. This was, for instance, done in the measurements by Baxter and the author on tungsten. Four or five volumes of silica had been mixed with one volume of tungsten trioxide powder prior to reduction. This method has been proved effective to avoid the formation of larger particles.

The angular range through which a certain reflection is scattered in so far as it is due to the limitation of the scattering crystallites can be taken as inversely proportional to the number of regularly spaced lattice planes*, while with crystallites extending uniformly in all directions their number in unit volume is inversely proportional to the third power of the linear dimensions. With powder particles oriented at random the number of particles which contribute to the scattering for any reflection for a given setting of the powder layer and a given aperture of the incident beam is thus inversely proportional to the fourth power of the linear size of the crystallites. This accounts for the rapid variation of this number with the size of the scattering units. The actual numbers depend on experimental conditions, and in particular on the volume from which scattering takes place and on the angular aperture of the beams. For average conditions the number of crystallites taking part in the scattering of a narrow incident beam with crystallites of 10^{-6} cm. is of the order of 10^7 to 10^8, which is more than adequate to give good averages. For a size of 10^{-5} cm. the number of scattering crystallites would be 10^3 to 10^4, while larger crystallites give distinct unevenness of the reflected intensities.

In order to improve the averaging, when reflections are recorded simultaneously without rotating screen and with the powder set at a certain angle α, we give a small rocking motion to the layer extending through approximately $0{\cdot}2^\circ$ to $0{\cdot}6^\circ$ about an axis vertical to the plane AOB, figure 1. This is particularly effective in reducing the unevenness due to the presence of a limited number of larger units. Such crystallites even when they represent only a small fraction of the total mass

* This refers to an arrangement with an approximately linear source.

and bring no appreciable extinction error, produce accidental inequalities which affect the photometric evaluation. For such units the range of reflection as determined from the dynamical theory of scattering is of the order of 5″. Thus a rocking motion of 30′ increases the number of reflecting elements with regard to any particular direction of incidence about 300 times. The rocking should be limited to a sufficiently small range, so as not to affect the intensity ratios as given by (8) and (8′). For large values of α and β the change of this expression with α for a small displacement to either side of the prefixed value is linear, so that the effects of symmetrical displacements to both sides cancel out. For small values of α and β this is not the case and the range of rocking should be made small accordingly. The treatment given to the deviations from the correct setting in an earlier paper[2] can be applied to this case.

§ 9. THE METHOD OF MIXED POWDERS

When the absolute value of F is being determined from the reflections of the powder layer we have to refer to expression (7) in the case of a flat layer or to a similar expression for other cases which require side by side with the measurement of the scattered intensity P_b the measurement of the incident beam and of the coefficient of absorption. It is difficult to introduce accurate numerical values for these quantities.

To meet these conditions we have resorted to the method of the quantitative powder mixture, which consists of using a powder layer containing a mixture of two kinds of crystals in the mass ratio $m_1 : m_2$[2]. If both types of particles are so small as to conform with the conditions of a uniform layer as discussed in §§ 3, 4 the same absorption coefficient applies to the reflections from each type of crystallites and we obtain P_b-values for each of them related to a common intensity of the incident beam and to a common absorption coefficient μ. The quantities N and μ do not therefore enter into the expression for relative values and we have for the ratio

$$\frac{P_{b(hkl)_1}}{P_{b(h'k'l')_2}} = \frac{m_1 \rho_2 p_{(hkl)} F^2_{(hkl)_1} N_1^2 q_1 \sin^2\theta_2}{m_2 \rho_1 p_{(h'k'l')} F^2_{(h'k'l')_2} N_2^2 q_2 \sin^2\theta_1} \cdot \frac{i_{(hkl)}}{i_{(h'k'l')}} \qquad \ldots\ldots(10),$$

where the suffixes 1 and 2 refer to the two kinds of particles and the suffixes (hkl) and $(h'k'l')$ mark two particular reflections.

The densities ρ_1 and ρ_2 of the two types of particles appear in this expression, as in equation (4), to account for the fact that the ratio of the masses per unit volume enter into P_b.

By this method the measurement of the ratio of the scattered intensities interlinks the values F_1 and F_2. An unknown F can be related to a value of F which is known from measurements on single crystals or for which calculated values are available.

§ 10. ABSORPTION EFFECTS IN COARSE POWDERS

The introduction of expression (10) implies that a common absorption term applies to both types of crystals. This is satisfied when the absorption in the individual particle produces a negligible reduction of intensity of the transmitted

X-ray beam and when the particles are sufficiently well mixed to conform with the requirement indicated in discussing the absorption coefficient of a powder in § 3.

These points have been overlooked in some cases and Schäfer[11] refers to the error which arises when the method of powder mixtures is applied to larger particles. He refers to this as accounting for discrepancies which appeared between measurements by Glocker and Schäfer[12] and the results of subsequent calculations by Hönl[13]. In discussing the error introduced by absorption he considers in the first place the effect of defective balancing of the absorption in a column of particles of a mixture containing a limited number of elements.

A different treatment may be given here, which considers in the first place the absorption in the individual particle and is applicable in the most direct way to the conditions of a powder layer.

In considering particles which possess appreciable absorption two distinct cases have to be distinguished. In the first case the particle consists of one single crystallite. In this case in general (but with the exception of very heavy atoms like uranium and very long X-ray wave-lengths of 2 A. or more) it can be shown by introducing numerical values that primary extinction begins to affect the measurements for a particle-size smaller than that for which absorption becomes appreciable. The requirement of negligible extinction thus sets a limit to the particle-size at which absorption can be neglected, so that no consideration need be given to an absorption effect in the individual particle. In the second case the individual crystallites are sufficiently small to conform to the condition that extinction and absorption be negligible, but the particles are formed of larger aggregates which may possess appreciable absorption. This is the case which needs more detailed consideration.

We assume, as the simplest model of a layer, particles of two microcrystalline substances 1 and 2 forming cubes of equal size arranged in strata parallel to the surface of the powder, as indicated by figure 3. Let z be the side of the cubes and

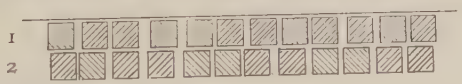


Figure 3.

μ_1 and μ_2 the absorption coefficients for the two types of particles. If we consider radiation incident and scattered in directions near to the normal to the surface so as to leave out oblique transitions from one particle into adjacent particles of different type, the scattering for the first stratum is given by the integration of equation (2) for a flat layer of thickness z. Thus we obtain

$$P_b = \frac{X p_{(hkl)} F^2_{(hkl)} N^2 q \chi}{8\pi b \sin\theta \mu (1+\sin\alpha/\sin\beta)} (1 - e^{-\mu z (1/\sin\alpha + 1/\sin\beta)}) \quad \ldots\ldots(11),$$

and for the ratio P_{b_1}/P_{b_2} of two reflexions (hkl) and $(h'k'l')$, changing as from (9) to (7)

$$\left(\frac{P_{b(hkl)_1}}{P_{b(h'k'l')_2}}\right)_{\mu_1\mu_2} = \frac{m_1 \rho_2 p_{(hkl)} F^2_{(hkl)_1} N_1^2 q_1 \sin^2\theta_2}{m_2 \rho_1 p_{(h'k'l')} F^2_{(h'k'l')_2} N_2^2 q_2 \sin^2\theta_1} \cdot \frac{\mu_2 \sin\beta_1 \cos\frac{1}{2}(\beta_2 - \alpha_2)}{\mu_1 \sin\beta_2 \cos\frac{1}{2}(\beta_1 - \alpha_1)} \cdot \frac{1 - e^{-\mu_1 z (1/\sin\alpha_1 + 1/\sin\beta_1)}}{1 - e^{-\mu_2 z (1/\sin\alpha_2 + 1/\sin\beta_2)}} \quad \ldots\ldots(12).$$

This value has to be compared with the value $(P_{b_1}\ P_{b_2})_\mu$ scattered from an intimately mixed layer, to which the coefficient of absorption as given by equation (4) applies, and containing the two constituents 1 and 2 in the same mass ratio. We assume the same scattering volume in each case and do not consider the absorption in the interspaces, which is not relevant.

R We obtain a measure for the absorption error R by comparing these ratios

$$R = \left(\frac{P_{b(hkl)_1}}{P_{b(h'k'l')_2}}\right)_{\mu_1\mu_2} \Big/ \left(\frac{P_{b(hkl)_1}}{P_{b(h'k'l')_2}}\right)_{\mu} \qquad \ldots\ldots(13).$$

Introducing the above values this becomes

$$R = \frac{\mu_2}{\mu_1} \cdot \frac{1 - e^{-\mu_1 z\,(1/\sin\alpha_1 + 1/\sin\beta_1)}}{1 - e^{-\mu_2 z\,(1/\sin\alpha_2 + 1/\sin\beta_2)}} \Big/ \frac{1 - e^{-\mu z\,(1/\sin\alpha_1 + 1/\sin\beta_1)}}{1 - e^{-\mu z\,(1/\sin\alpha_2 + 1/\sin\beta_2)}} \qquad \ldots\ldots(13').$$

R expresses how much the ratio of the intensities $(P_{b_1}\ P_{b_2})_\mu$ scattered from the first stratum is modified by the fact that the two types of particles present different absorption compared with the ratio of these intensities if the same absorption coefficient applied to both. If $\mu_2 > \mu_1$, the term R is the factor by which the intensity P_{b_2} should be multiplied to correct for the greater absorption loss.

The correction is definite when the directions of incidence and scattering are normal to the layer. Considering this case more in detail when $\alpha - \beta = 90$ the expression simplifies thus

$$R = \frac{\mu_2}{\mu_1} \Big/ \frac{1 - e^{-2\mu_1 z}}{1 - e^{-2\mu_2 z}} \qquad \ldots\ldots(14).$$

We see that this becomes unity when $\mu_2 = \mu_1$.

When $e^{-2\mu_1 z}$ and $e^{-2\mu_2 z}$ are both small, i.e. when the absorption in the individual particle is large, expression (14) becomes μ_2/μ_1, which indicates the maximum error with coarse particles. When the absorption in both types of particles is small, i.e. when $e^{-2\mu_1 z}$ and $e^{-2\mu_2 z}$ are both nearly unity, the limiting value of expression (14) becomes equal to unity, as we should expect.

Table 2. Term (14) for scattering in stratum I for various ratios of absorption coefficient μ_2/μ_1 and various values of absorption. The absorption is expressed by the intensity loss $1 - e^{-\mu_2 z}$ for transmission in the more absorbing particle.

μ_2/μ_1 \ $1 - e^{-\mu_2 z}$	0·05	0·1	0·2	0·3	0·4	0·5	0·7	0·9
1·5	1·019	1·034	1·072	1·113	1·158	1·208	1·318	1·442
2	1·022	1·052	1·111	1·178	1·249	1·334	1·540	1·818
4	1·047	1·078	1·170	1·281	1·410	1·560	2·405	2·760
10	1·058	1·103	1·223	1·350	1·520	1·727	2·350	3·725

Intermediate values are illustrated in table 2, in which are plotted values of expression (14) for different values μ_2/μ_1 and different values of $1 - e^{-\mu_2 z}$, which quantity measures the absorption of radiation in passing through the particle with greater μ. With great disparity between μ_2 and μ_1 the error is appreciable even

when the individual particles absorb little. Thus when $\mu_2/\mu_1 = 10$ with an intensity-loss of only 5 per cent in the more absorbing particle the error is 5·8 per cent. When $\mu_2/\mu_1 = 1{\cdot}5$ with a corresponding loss of 10 per cent the error is only 3·4 per cent.

Successive strata give decreasing contributions to the scattered intensity owing to the absorption in the strata in front. When the absorption in the front stratum is very large, the successive layers will not alter appreciably the intensity-ratios as scattered by the front stratum and the discrepancy as compared with the intensity-ratio obtained from a finely subdivided powder is given by the ratio μ_2/μ_1 of the absorption factors. Such a case is of no practical interest.

In considering successive strata with particles having little absorption in each individual stratum, we have to distinguish the case in which corresponding particles in adjacent layers are of the same type figure 3 (*a*), and that in which they are of different type, figure 3 (*b*). In the first case the absorption error introduced by the size of the particles will be increased owing to the exponential character of the absorption law. In carrying out any closer estimates it will be necessary to take account of the shape- and size-distribution of the particles of each kind and of their numbers. It can however be seen that for the scattering from a greater number of strata the increased absorption effect due to the superposition of a large number of particles of the same type decreases. The probabilities of definite distributions are determined by terms in which the number n of the elements enters as a factorial, n
while the absorption is determined by terms in which n enters as an exponential. With increasing n the factorial terms gain over the exponential terms, which means that irregularities in the grouping of the particles become of lesser significance in determining the intensities. When we consider the case of oblique incidence the effect of the incident or the scattered beam passing through adjacent blocks can be referred to the case of transition through successive strata. Any more detailed treatment depends on the introduction of definite assumptions as to the shape and size of the particles and as to their relative numbers for which no actual data are available. The schematic treatment shows however that the absorption error of large particles as given by table 2, for one stratum, offers a criterion for estimating with a certain margin of safety the magnitude of the possible absorption error for a thicker layer.*

The practical question as to how far particles consisting of large aggregates of crystallites can actually be used for intensity-measurements depends not only on the absorption effect we have discussed, but also on the extent to which the particles

* When the layer consists of particles which satisfy the condition of negligible absorption in the individual particle and which are intimately mixed, the chance must be considered of several particles of the same type being placed close together when we obtain what is equivalent to an aggregate. Again, it is not possible to evaluate the irregularities introduced by such groups without introducing specialized assumptions as to the distribution with regard to sizes and numbers of particles. We can, however, obtain some estimate of the differential absorption effect which may arise from such groupings, by referring to expression (14) and introducing into it a small multiple of the mean particle-size. Large chance irregularities are very rare and chance groupings resulting from particles at larger distances being placed in line for any particular direction can be neglected when the layer is rotated or rocked.

can be considered to be free of larger crystallites. We have observed not only that in the case of ductile metals like gold, silver and platinum does the tendency exist for microcrystals to unite to larger crystallites, but also that in the case of brittle metals with a high temperature of recrystallization like tungsten some larger crystals are found embodied in the microcrystalline powder when no steps are taken to avoid the formation of large particles; the same thing was found with binary compounds like rock-salt. In such cases the intensity-measurements are affected by extinction. We consider therefore that in general, so far as the presence of larger crystallites is not eliminated, extinction effects are a potential cause of error and that with aggregate particles it is not safe to rely on their subdivision into uniformly small micro-crystallites. The best way for avoiding these errors is to limit the maximum extension of the crystallites by the actual size of the particles to a value for which extinction is small. In this case, as has been indicated above, the absorption error in the individual particle is negligible.

§ 11. ACCURACY OF INTENSITY-MEASUREMENTS WITH POWDERS

As to the accuracy obtainable with these methods, it must be considered that when reflections are compared by means of expressions depending on the intensity relation (8) the intensities depend on introducing the correct angular values. For large angles α and β the expressions are not sensitive to small angular deviations, so that the accuracy with which the angular setting can be measured scarcely affects the results; when one of the angles is small the limits of error within which the correct setting can be established and in particular the limits of trueness of the powder surface make themselves felt and the error is found to be about 2 per cent for glancing angles of 15°, but increases rapidly for smaller angles.

A certain intrinsic error is attached to the photographic photometry. This is mainly expressed by the deviations found when repeated measurements of the same group of lines taken under similar conditions are compared. The limits of error introduced from this source with measurements of suitable lines, i.e. lines representing a well defined intensity-step, showing a flattened peak with rapid fall of density to both sides, is about 1 per cent, which is mainly determined by the local inequalities of the photographic emulsion. When lines with a gradual decrease of density are measured the accuracy is less. With measurements of the reflections from platinum black, which gives broadened lines, an accuracy of 6 per cent was found and the accuracy found with lines situated at greater distance from the focused angle, like the broader lines in figure 2, is similar. The error depends then very much on the distribution of the density through the line. With lines corresponding to a vanishing size-distribution much larger errors are observed. Lines of this type are unsuitable for quantitative measurements.

REFERENCES

(1) Brentano, J. C. M. and Baxter, A. *Z. Phys.* **89**, 720 (1934).
(2) Brentano, J. C. M. *Phil. Mag.* **4**, 620 (1927); **6**, 178 (1928).

(3) Brindley, G. W. and Spiers, F. W. *Proc. phys. Soc.* **46**, 841 (1934).
(4) Darwin, C. G. *Phil. Mag.* **44**, 433 (1926). Bragg, W. L., Darwin, C. G. and James, R. W. *Phil. Mag.* **1**, 897 (1926).
(5) Ewald, P. *Handb. der Physik*, **23** (2), (1933).
(6) Claassen, A. *Phil. Mag.* **9**, 57 (1930).
(7) Rusterholz, A. *Helv. phys. Acta*, **4**, No. 2 (1931).
(8) Brentano, J. C. M. *Proc. phys. Soc.* **37**, 184 (1925).
(9) Brentano, J. C. M. and Adamson, J. *Phil. Mag.* **7**, 507 (1929).
(10) Brentano, J. C. M. *Z. Phys.* **70**, 74 (1931).
(11) Schäfer, K. *Z. Phys.* **86**, 738 (1933).
(12) Glocker, R. and Schäfer, K. *Z. Phys.* **73**, 289 (1933).
(13) Hönl, H. *Z. Phys.* **84**, 1 (1933); *Ann. Phys.*, Lpz., (5) **18**, 625 (1933).
(14) Brentano, J. C. M., Baxter, A. and Cotton, F. W. *Phil. Mag.* **17**, 370 (1934).

DISCUSSION

Messrs G. W. Brindley and F. W. Spiers. In reading this paper our attention has been directed to a previous paper by the author in the *Zeitschrift für Physik* **70** (1931) which had escaped our notice. In our previous paper[3] we discussed and gave references to the author's previous work and very much regret not having seen his 1931 paper in which the use of a flat stationary plate of powder was described; we offer him our sincere apologies. In putting forward our own paper we wished to emphasize particularly the method of controlling the intensities of X-ray lines by the variable focusing and absorption obtained with a flat layer of powder, a method to which the author raises certain objections in § 1 of his paper. In our own work we have found the method simple to use and reliable.

Author's reply. I regret if my note conveys the impression of a too strong criticism of the paper by Messrs Brindley and Spiers. Not only in the 1931 paper, but also in later work, as in that of Baxter and the author where use was made of a stationary or quasi-stationary layer, the lines were of necessity not equally in focus and the weaker lines were then placed nearer to the focused angle so that their peak values were enhanced. We found it better, however, to confine ourselves to a small angular range and to observe the lines near to the focused angle using at the same time a wide entrance slit, in preference to placing the lines far outside the focused range, where the change of the i factor becomes important enough to constitute actually a controlling factor of the intensity ratios. In the present paper, the object of which is to set forth the conditions for obtaining the best accuracy with powder methods, these various aspects had to be discussed. As has been indicated in the paper certain conditions are more suitable for greater angular ranges than others; they are found in particular with light atoms and large deflections, where the loss of accuracy in comparing lines at a great distance from the focused angle is less marked.

535·421

THE THEORY OF THE FORMATION OF AN IMAGE BY A PLANE BAND GRATING USED IN THE SOFT X-RAY REGION

By F. F. P. BISACRE, O.B.E., M.A., B.Sc.

Received April 29, 1935

ABSTRACT. This paper describes an investigation of the reflection of light from a plane grating illuminated by a cylindrical wave diverging from a Huyghens line source. It is shown that for pencils sufficiently narrow the isophasics (properly parabolas) become circles with a virtual focus as centre. The position of this virtual focus is found. A formula is given for the intensity of the light along the reflected isophasic. It has a Fresnel integral as a factor. It is shown that if the dimensions of the grating are chosen so that the modulus of the Fresnel integral has its maximum value, the intensity along the isophasic varies after the fashion of an Airy image and the maximum possible intensity occurs on the axis of the reflected pencil—in other words, the reflected light is automatically focused. The critical (optimum) length of the grating for automatic focusing is determined by the condition that the quadratic term in the expansion for the optical path in powers of distance measured along the grating face from its centre must be three-eighths of a wave-length. Formulae for the dispersion and resolution of an optimum grating are given and the resolution turns out to be exactly the same as for an ordinary grating, namely the total number of lines on the grating multiplied by the spectral order. Some numerical examples are given. Third-order effects have been considered and it is shown that in the conditions contemplated in the use of these optimum gratings, the third-order term affects the length of the optical path by something like one 650th part of a wave-length, and is consequently negligible.

For some time past plane gratings have been successfully used for photographing soft X-ray spectra, from say 10 to 75 A., without lenses. These gratings are lightly ruled on glass and are used at grazing angles of incidence. Glass is chosen instead of a metal because of its high resistance to corrosion[1] in soft X radiation. The grating operates by means of the flats between the grooves and is lightly ruled to keep the width of each flat roughly half the pitch—the best proportion. It must be used at grazing angles to ensure a high reflecting-power[2]. The gratings are ruled with a band a few hundred lines only, instead of the usual several thousand. The precise width of the band is a point of considerable importance. That such a grating if ruled with the right number of lines does possess remarkable automatic focusing power—much as a concave grating does—is proved by the excellent photographs of spectra reproduced in papers by Siegbahn and Magnusson[3] and J. A. Prins[4] for instance.

Prins gives a rough rule for determining the proper number of lines: The quadratic term in the expansion for the optical path in powers of distance measured along the grating face from its centre must not exceed half a wave-length. It may

of interest to examine the theory of this automatic focusing in some detail, in the pe of obtaining definite information as to (*a*) the optimum width of the band of es; (*b*) the nature of the images formed—in particular the nature of the isophasics d the distribution of intensity along them; and (*c*) the positions of the centres of se images and consequently the validity of the usual formula giving the directions the lines.

Siegbahn and Magnusson's apparatus[3] is sketched in figure 1. The whole paratus is contained in a vacuum chamber which can be attached to a vacuum mp. The X-ray generating apparatus is on the left. The target is made of a osen material, such as tungsten, and the electrical pressure used was some oo V. Two slits are provided, the first being an ordinary mechanical slit some

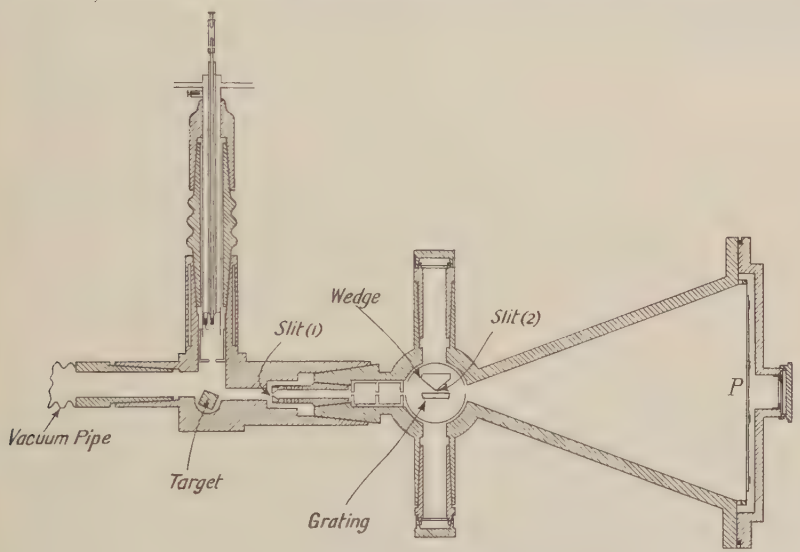


Figure 1. Vacuum plane-grating spectrometer*.

01 mm. wide while the second is formed by a wedge placed very close to the ce of the grating, on the plane originally used by Seemann in his crystal grating ork. The photographic plate is at P and is at right angles to the central axis of the ts. The back of the grating is not parallel to the front face on which the lines are led, so that light reflected from the back of the grating is thrown into what is ually a harmless direction. The bands of lines used by Siegbahn on different atings ranged from 1 to 3 mm. He built a special ruling-machine and ruled a set gratings with from 300 to 1800 lines per mm. and bands of 1 to 3 mm. The best sults appeared to be got with a 1-mm. band having 1800 lines, a slit-width S_1 of out 0·01 mm. and an incidence grazing angle of about 1°. The paper does not ate the wedge slit-width; glancing angles of incidence from 0·5° to 3° are mentioned.

The use of two parallel slits with an extended source of radiation secures (*a*) a rrow band of the source as the illuminant; (*b*) narrow pencils of light from the

* From Siegbahn and Magnusson, *Z. Phys.* **62**, 443 (1930), by the courtesy of Julius Springer, rlin.

elementary Huyghens sources constituting the active source-band; and (*c*) that th axes of these pencils shall be all very nearly coaxial with the centre-line of the slit

The angular aperture of the pencils is probably less than 1 minute of arc wit the actual dimensions of the apparatus and slit-widths used. At all events, they a of this order.

We are thus led directly to a study of the screen image formed when a ver narrow pencil of light from a Huyghens source is reflected from a plane gratin The actual image will be the sum of these elementary (non-coherent) images comi from the Huyghens sources forming the exposed active source-band, as defined b the pair of slits. We assume that the angular aperture of the incident Huyghe pencil, as controlled by the slits, is sufficiently wide to overlap the grating slightl so that any diffraction ripples on the edges of the incident wave-front do not actual reach the grating.

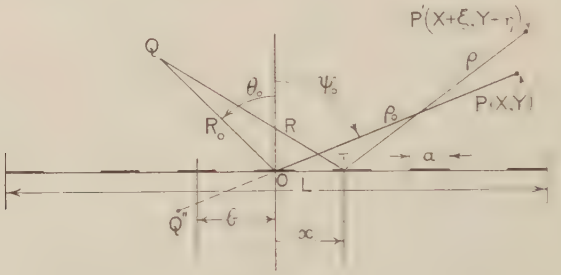


Figure 2.

We shall treat the problem as a 2-dimensional one. Let Q, figure 2, be a Huygh

X, Y, ξ, η source, P a point (X, Y), and P' a neighbouring point $(X+\xi, Y-\eta)$. Let O the centre of the grating, assumed to be at the centre of a reflecting strip.

R_0, θ_0 Let R_0, θ_0 define the position of Q in polar coordinates, the positive direct

ρ_0, ψ_0 of the angle being as shown in the figure; and let ρ_0, ψ_0 define the position of again with ψ_0 positive in the direction of the arrow in the figure.

The coordinates ξ, η fix P' relative to a set of rectangular axes through P a parallel to the main rectangular axes, viz. the grating-face and the normal at O, shown in the figure.

L The grating extends a distance $\frac{1}{2}L$ on each side of O.

x Let T be any point on one of the reflecting strips whose coordinates are x

R, ρ Let R be equal to QT and ρ to $P'T$.

We suppose the light to be polarized in the plane of incidence, so that

E_z' incident electric vector E_z' is wholly along the Oz axis, i.e. it is perpendicular to plane of the figure. It is given by

$$E_z' \approx \frac{B}{\pi} \cdot e^{i\pi/4} \cdot \frac{e^{-i\alpha R}}{\sqrt{R}} \cdot e^{ipt} \qquad \text{.....(1)}$$

B where B is a constant, this being a convenient asymptotic form for the exact Be function representing a divergent cylindrical wave from a line source. The quan

α, λ, p α is $2\pi/\lambda$ where λ is the wave-length, and p is 2π times the frequency of the lig

It can be shown that the value of the electric vector $E_{z,P'}$ of the reflected field at P' is given by the integral, taken over the reflecting surface of the grating,

$$E''_{z,P'} \approx -\frac{iB}{\pi\sqrt{\lambda}} \int \frac{e^{-i\alpha(R+\rho)}}{\sqrt{R\rho}} . \cos nR . ds . e^{ipt} \qquad \ldots\ldots(2).$$

$E''_{z,P'}$

provided the reflecting power of the surface is unity. Consequently this is a suitable limiting form for both metals and glass if these are used under conditions where the reflecting-power is approximately unity. If this condition is not satisfied no such formula as equation (2) is available and the problem becomes much more complex. We can therefore take equation (2) as a guide, at all events, to what happens provided the reflecting surface has a reflecting-power sufficiently near to unity.

When $E''_{z,P'}$ has been determined, the whole reflected electromagnetic field follows at once. A simple calculation shows that

$$R = R_0 + x \sin\theta_0 + x^2 . \frac{\cos^2\theta_0}{2R_0} + \text{higher terms in } x.$$

If the incident light were parallel the expression would end exactly with $x \sin\theta_0$.

Also $$\rho^2 = (X+\xi-x)^2 + (Y+\eta)^2.$$

$$\therefore \quad \rho \approx \rho_0 + (\xi - x)\sin\psi_0 + \eta\cos\psi_0 + \frac{1}{2\rho_0}[(\xi-x)\cos\psi_0 - \eta\sin\psi_0]^2 + \ldots,$$

where $$X = \rho_0 \sin\psi_0 \text{ and } Y = \rho_0 \cos\psi_0.$$

Hence $$R+\rho \approx R_0 + \rho_0 + x(\sin\theta_0 - \sin\psi_0) + \xi\sin\psi_0 + \eta\cos\psi_0$$
$$+ \frac{1}{2\rho_0}[(\xi-x)\cos\psi_0 - \eta\sin\psi_0]^2 + \frac{1}{2R_0}x^2\cos^2\theta_0 + \ldots$$

as far as the second order in x, ξ and η.

Now $R+\rho$ occurs in our integral as

$$e^{-i\alpha(R+\rho)} = e^{-i(2\pi/\lambda)(R+\rho)},$$

so it is only the excess of $(R+\rho)$ over *any* multiple of λ that determines the value of this expression.

Let us put x equal to $nb+\bar{x}$, where n is a positive or negative integer, b is the grating-pitch and $|\bar{x}| \leqslant a/2$, where a is the width of a flat strip. n, b $|\bar{x}|, a$

Also put q equal to $(\sin\theta_0 - \sin\psi_0)$. Then q

$$qx = qnb + q\bar{x}$$

and, if $$qb = r\lambda, \text{ where } r = \pm 1, 2, 3, \ldots \qquad r$$

$$qx = (\text{a multiple of } \lambda) + q\bar{x}.$$

We can therefore replace qx in the expression for $R+\rho$ by $q\bar{x}$, where

$$-\frac{a}{2} \leqslant \bar{x} \leqslant \frac{a}{2},$$

and $\bar{x}$ is now comparable to λ whereas x itself is not.

But the condition $qb=r\lambda$ is the ordinary grating formula

$$b\,(\sin\theta_0-\sin\psi_0)=r\lambda,\ r=\pm 1,\ 2,\ 3,\ \ldots.$$

The integral in equation (2) becomes, approximately,

$$I\approx\frac{\cos\theta_0}{\sqrt{(R_0\rho_0)}}\,e^{-i\alpha(R_0+\rho_0)}\,.\int e^{-i\alpha q\bar{x}}\,.\,e^{-i\alpha P}\,a\bar{x},$$

where

P $$P=(\xi\sin\psi_0+\eta\cos\psi_0)+\frac{1}{2\rho_0}\,[(\xi-x)\cos\psi_0-\eta\sin\psi_0]^2+\frac{1}{2R_0}\cos^2\theta_0x^2+\ldots$$

$$=(\xi\sin\psi_0+\eta\cos\psi_0)+A\left[x-\frac{(\xi\cos\psi_0-\eta\sin\psi_0)\cos\psi_0}{2A\rho_0}\right]^2$$

$$-\frac{\cos^2\psi_0\,(\xi\cos\psi_0-\eta\sin\psi_0)^2}{4A\rho_0^2}+\frac{(\xi\cos\psi_0-\eta\sin\psi_0)^2}{2\rho_0}+\ldots$$

A and $$A=\frac{1}{2}\left\{\frac{\cos^2\psi_0}{\rho_0}+\frac{\cos^2\theta_0}{R_0}\right\} \qquad \ldots\ldots(3).$$

Now choose ξ and η so that

$$(\xi\sin\psi_0+\eta\cos\psi_0)+\frac{(\xi\cos\psi_0-\eta\sin\psi_0)^2}{2\rho_0}\left\{1-\frac{\cos^2\psi_0}{2A\rho_0}\right\}=0.$$

The term $\{1-(\cos^2\psi_0/2A\rho_0)\}$ is $\dfrac{\rho_0\cos^2\theta_0}{R_0}\Big/\left(\cos^2\psi_0+\dfrac{\rho_0}{R_0}\cos^2\theta_0\right)$

and the equation becomes

$$2\,(\xi\sin\psi_0+\eta\cos\psi_0)+\frac{(\xi\cos\psi_0-\eta\sin\psi_0)^2\cos^2\theta_0}{R_0\cos^2\psi_0+\rho_0\cos^2\theta_0}=0.$$

ξ', η', ψ_0 Now change the axes at P from ξ, η to ξ', η', where the $\angle\,\xi' P\xi=\psi_0$, figure 3.

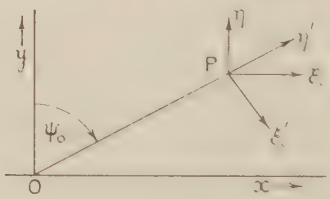


Figure 3.

Then

$$\xi'=\xi\cos\psi_0-\eta\sin\psi_0,$$

$$\eta'=\xi\sin\psi_0+\eta\cos\psi_0,$$

h so, if $\{R_0\cos^2\psi_0+\rho_0\cos^2\theta_0\}/\cos^2\theta_0=h$, the equation becomes

$$\xi'^2+2h\eta'=0 \qquad \ldots\ldots(4),$$

which is a parabola, concave towards O.

So far as the radius of curvature at P is concerned this is indistinguishable from the circle

$$\xi'^2+\eta'^2+2h\eta'=0.$$

The centre of curvature is therefore at Q'' ($\xi'=0$, $\eta'=-h$) i.e., is on PO produced, where

$$OQ''=R_0\cos^2\psi_0/\cos^2\theta_0 \qquad \ldots\ldots(5).$$

Along the locus defined by equation (4) we have, then, for P P

$$P = A\left[x - \frac{\xi' \cos\psi_0}{2A\rho_0}\right]^2$$

$$= At^2, \text{ say.} \qquad t$$

Hence $$\int e^{-i\alpha q\bar{x}} . e^{-i\alpha P}\, dx = \int_{-a/2}^{+a/2} \sum_{n=-m}^{+m} . e^{-i\alpha(q\bar{x}+At^2)}\, d\bar{x}, \qquad n$$

where $$t^2 = \left[(nb+\bar{x}) - \frac{\xi' \cos\psi_0}{2A\rho_0}\right]^2.$$

Since n enters into t^2, only, we can put the integral into the form

$$\int_{-a/2}^{+a/2} e^{-i\alpha q\bar{x}} . \left\{ \sum_{n=-m}^{+m} . e^{-i\alpha A t^2} \right\} d\bar{x}.$$

The sum is the sum of the values of the function

$$e^{-i\alpha A[X-C]^2},$$

where $$C = \xi' \cos\psi_0 / 2A\rho_0,$$

at $$X = \bar{x} - mb,\ \bar{x} - \overline{m-1}b,\ \ldots\ \bar{x} + mb,$$

all at distances b apart.

These are close enough (as examination in detail shows) to allow us to replace the sum by the integral,

$$\frac{1}{b}\int_{\bar{x}-mb}^{\bar{x}+mb} e^{-i\alpha A[X-C]^2}\, dX = \frac{1}{b} \cdot \sqrt{\left(\frac{\pi}{2\alpha A}\right)} \cdot \int_j^k e^{-i\frac{\pi}{2}u^2}\, du$$

where

$$j = \sqrt{\left(\frac{2\alpha A}{\pi}\right)}(\bar{x} - mb - C)$$

$$k = \sqrt{\left(\frac{2\alpha A}{\pi}\right)}(\bar{x} + mb - C) \qquad \ldots\ldots(6).$$

and $$C = \xi' \cos\psi_0 / 2A\rho_0$$

We can therefore put the sum as

$$\frac{1}{b}\sqrt{\left(\frac{\pi}{2\alpha A}\right)} . F(\bar{x}), \text{ where}$$

$F(\bar{x})$ is a limited Fresnel integral from j to k and involving $\bar{x}$ through its limits. $F(\bar{x})$

Consequently we have

$$E_{z,P''} \approx -\frac{iB}{\pi\sqrt{\lambda}} \cdot \frac{\cos\theta_0}{\sqrt{(R_0\rho_0)}}\, e^{-i\alpha(R_0+\rho_0)} \cdot \frac{1}{b} \cdot \sqrt{\frac{\pi}{2\alpha A}} \cdot \int_{-a/2}^{+a/2} e^{-i\alpha q\bar{x}} . F(\bar{x})\, d\bar{x} \qquad \ldots(7).$$

This result of course includes all diffraction effects due to the limitation in the range of the Fresnel integral. It is only in quite exceptional cases that the presence of $\bar{x}$ in j and k affects the value of the Fresnel integral appreciably, for $\bar{x}$ is a very small quantity comparable to λ, whereas j and k are comparable to the grating-length.

As a rule, then, $F(\bar{x})$ can be taken outside the integration sign.

We then have $$\int_{-a/2}^{+a/2} e^{-i\alpha q\bar{x}}\, d\bar{x} = \frac{a \sin\left(\frac{\pi r a}{b}\right)}{\left(\frac{\pi r a}{b}\right)},$$

so that $$E_{z,P''} \approx -\frac{iB}{\pi\sqrt{\lambda}} \cdot \frac{\cos\theta_0}{\sqrt{(R_0\rho_0)}} \cdot e^{-i\alpha(R_0+\rho_0)} \cdot \sqrt{\frac{\pi}{2\alpha A}} \cdot \frac{a}{b} \cdot F \cdot \frac{\sin\frac{\pi r a}{b}}{\frac{\pi r a}{b}},$$

where
$$\left.\begin{aligned} F &= \int_j^k e^{-i\frac{\pi}{2}u^2}\, du \\ j &= -\sqrt{\frac{2\alpha A}{\pi}}\left(\frac{L}{2}+C\right) \\ k &= +\sqrt{\frac{2\alpha A}{\pi}}\left(\frac{L}{2}-C\right) \\ C &= \xi' \cos\psi_0 / 2A\rho_0 \end{aligned}\right\} \quad \text{.....(8).}$$

j *k* *C* and

This formula shows that the phase of $E_{z,P''}$ is constant, except in so far as it is modified by the single factor F, the appropriate Fresnel integral.

We have localized P' so that it lies on the parabola through P, defined by equation (4) and having its centre of curvature at Q'', whose distance from O is $R_0 \cos^2\psi_0/\cos^2\theta_0$, by equation (5).

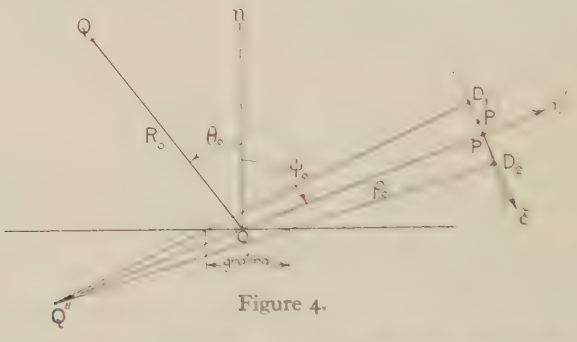


Figure 4.

This parabola is therefore an isophasic in the same sense as that in which a straight line is an isophasic for a fragment of a plane wave; it is an isophasic except for phase-ripples coming from varying values of the Fresnel integral, i.e. except for the diffraction that accompanies every wave-fragment at its edges. We have thus determined the basic isophasic for the reflected field, the central point of which is fixed by the grating formula

$$b(\sin\theta_0 - \sin\psi_0) = r\lambda, \quad r = \pm 1, 2, 3 \qquad \text{.....(8a).}$$

Special case. The field at P itself can be easily got. At P, $\xi' = \eta' = 0$, hence $C = 0$ and the range of F is from

$$-\sqrt{\frac{2\alpha A}{\pi}} \cdot \frac{L}{2} \text{ to } +\sqrt{\frac{2\alpha A}{\pi}} \cdot \frac{L}{2} \qquad \text{.....(8b).}$$

If each of these limits is sufficiently large to enable us to replace them, in the Fresnel integral, by $\mp\infty$, we get $F=\sqrt{2}.e^{-i\pi/4}$ and

$$E_{z,P}''\approx-\frac{Be^{i\pi/4}}{\pi}\cdot\frac{\cos\theta_0}{\sqrt{(R_0\cos^2\psi_0+\rho_0\cos^2\theta_0)}}\cdot\frac{a}{b}\cdot\frac{\sin\dfrac{\pi ra}{b}}{\dfrac{\pi ra}{b}}\cdot e^{-i\alpha(R_0+\rho_0)} \quad \text{.....(9)},$$

the time factor being omitted.

When $r=0$, $\psi_0=\theta_0$ and if $a/b=1$,

i.e., when the grating is unruled, we get at once

$$E_{z,P}''\approx-\frac{Be^{i\pi/4}}{\pi}\cdot\frac{1}{\sqrt{(R_0+\rho_0)}}\,e^{-i\alpha(R_0+\rho_0)}.e^{ipt} \quad \text{.....(10)},$$

a result agreeing exactly with equation (1). The source is now at the reflected image, Q' in the mirror, and there is a reversal of phase of π at reflection. Our formula (8) is therefore confirmed in this simple limiting case.

This result, at the centre of the image, applies only if the mirror is long enough to give us asymptotic Fresnel integrals.

To draw the isophasic, we find ψ_0 from the formula

$$b\,(\sin\theta_0-\sin\psi_0)=r\lambda,\ r=\pm1,\ 2,\ 3,\ \ldots \quad \text{.....(11)}$$

and set off OP, figure 4, equal to the given value ρ_0.

Produce PO backwards to Q'' where

$$OQ''=R_0\cos^2\psi_0/\cos^2\theta_0.$$

Q'' is the virtual image. Rule lines $Q''D_1$ and $Q''D_2$ through the ends of the grating, cutting the parabola defined by equation (4), in D_1 and D_2. Then D_1PD_2 is the fragment of the isophasic that interests us and we propose now to discuss the diffraction effects at the edges of this fragment.

Now $k-j=\sqrt{(2\alpha A/\pi)}\,L$ hence as we give C in equation (8) an increasing value, we merely decrease the upper limit and the lower limit. We therefore throw the arc of the Cornu spiral (of constant length $k-j$) towards one or other of the asymptotic points. If L is sufficiently large for $(k-j)/2$ to give approximately the asymptotic value of the integral at the centre of the image ($C=0$), then as we proceed along the isophasic from the central point P there will be little change in modulus or phase until j (or k) is zero, i.e., from equation (8),

$$\xi'=A\rho_0L\sec\psi_0 \quad \text{.....(12)}.$$

when the amplitude will be halved.

Now the length of the arc D_1D_2, figure 4, cut off by the rays from Q'' to the ends of the grating is easily seen to be given by

$$D_1D_2\approx2A\rho_0\sec\psi_0.L \quad \text{.....(13)};$$

hence the breadth of the image (defined as the width between half-maximum values of the amplitude) is the geometrical image.

We shall thus get the usual rectangular image flanked by ripples of phase and amplitude as sketched in figure 5, and the film will record a broad band of width D_1D_2.

These remarks apply when the image is received on a film normal to the beam. If the film is perpendicular to the plane of the grating, then the pencil is cut obliquely by the film and this obliquity may have to be allowed for. This image shows no focusing property at all. The light is merely spread out over a band which is the projection of the grating from the virtual focus Q'' on to the film.

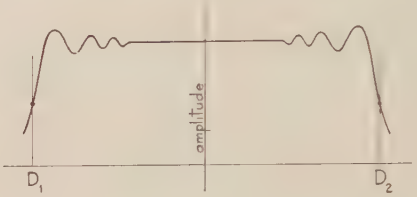


Figure 5.

Optimum (or focusing) conditions. Suppose now that instead of the upper limit of the integral $\int_0^d e^{-\frac{1}{2}\pi u^2}\,du$ being large enough to give asymptotic values, it is chosen just large enough to give the maximum modular value of this integral (not its asymptotic value). The upper limit will now have quite a small and perfectly definite value, viz., $d=\sqrt{\frac{3}{2}}$*, and

$$\int_{-d}^{+d} e^{-i\frac{1}{2}\pi u^2}\,du = 2\times 0{\cdot}945\; e^{-i\pi/4}$$

instead of

$$\int_{-\infty}^{+\infty} e^{-i\frac{1}{2}\pi u^2}\,du = 2\times 0{\cdot}707\; e^{-i\pi/4},$$

so that the central amplitude is now increased by 34 per cent and the intensity by 80 per cent.

If then we limit the length of the grating to secure this condition, we shall materially increase the central intensity of the image and modify the whole structure of the image; in fact, as we shall now show, we get an Airy type of image, instead of the rectangular pattern.

To secure this critical optimum condition, we get at once from equation (8*b*),

$$\sqrt{\frac{2\alpha A}{\pi}}\cdot\frac{L}{2}=\sqrt{\frac{3}{2}}$$

or, since

$$\alpha=2\pi/\lambda,$$

$$A.\left(\frac{L}{2}\right)^2=\frac{3}{8}\lambda \qquad \text{.....(13}a\text{),}$$

where

$$A=\frac{1}{2}\left(\frac{\cos^2\theta_0}{R_0}+\frac{\cos^2\psi_0}{\rho_0}\right)$$

$=$ the coefficient of the second-order term in the expansion of $(R+\rho)$ for the central image point P, figure 2, in powers of x.

* This is not *exactly* the value of a but it is within 1 per cent of the true value and is used in this form for arithmetical convenience.

Hence $A.(L/2)^2$ is the maximum value of this second-order term and we get the rule: The extreme value of the second-order term, in the expansion of $R+\rho$ in powers of x, measured along the grating face from its centre, should be exactly three-eighths of the wave-length. If it is either more or less than this value, the central intensity and sharpness of the image will be impaired. Prins has given this critical value as $\lambda/2$, which appears to be rather on the high side.

In figure 6 the dark line shows the range of integration for the point at the centre of the image. As we move from the centre of the image outwards the arc of integration AB moves in one or other direction, remaining constant in length. Suppose it moves in the direction of the arrows in figure 6.

Then, clearly from figure 6, oscillations of modulus and phase in F begin immediately, so that changes of amplitude and phase in the image begin at once. In other words, diffraction effects begin at once.

The effect is to give an amplitude-band as illustrated roughly in figure 7, while the phase-ripple begins, very feebly, at the centre of the image. This type of image clearly does show a focusing effect. The range of integration is $2\sqrt{\frac{3}{2}}$.

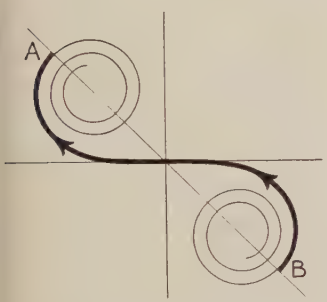


Figure 6.

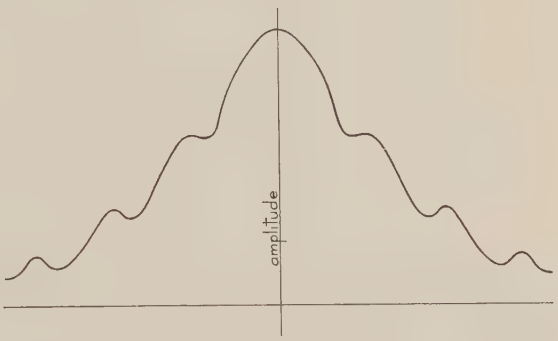


Figure 7.

The amplitude will, then, be proportional to y, where

$$y=\left|\int_{x-2\sqrt{(3/2)}}^{x} e^{-i\frac{1}{2}\pi u^2}\,du\right| \qquad \text{.....(14),}$$

where $x=\sqrt{\frac{3}{2}}$ at the centre of the band.

Differentiating the integral in equation (14) with respect to x and equating to zero, we get turning-values of y (but not all of them), where

$$x=\sqrt{\tfrac{3}{2}}+\sqrt{\tfrac{2}{3}}.n,\ n=1,\ 2,\ 3,\ \ldots \qquad \text{.....(15).}$$

The first occurs at $n=0$, i.e., at $x=\sqrt{\frac{3}{2}}$, the central line of the image.

Since $x_n=2\sqrt{\dfrac{A}{\lambda}.\left(\dfrac{L}{2}-C_n\right)}$, we get

$$\xi_n'=\pm A\rho_0 \sec\psi_0\sqrt{\frac{2\lambda}{3A}.n} \qquad \text{.....(16),}$$

But if L is the optimum length,

$$AL^2/\lambda=\tfrac{3}{2},\ \text{by (8}b\text{)} \qquad \text{.....(16}a\text{),}$$

$$\therefore\ \xi_n'=\pm\rho_0\lambda n/L\cos\psi_0 \qquad \text{.....(17).}$$

V If we denote by $2V$ the angle subtended by the grating at the centre of the image,

$$L \cos \psi_0/\rho_0 = 2V,$$

$$\therefore \quad \xi_1' = \lambda/2V \qquad \text{.....(17}a\text{)},$$

which is precisely the same as the half-width of the Airy image formed by a lens of angular aperture $2V$.

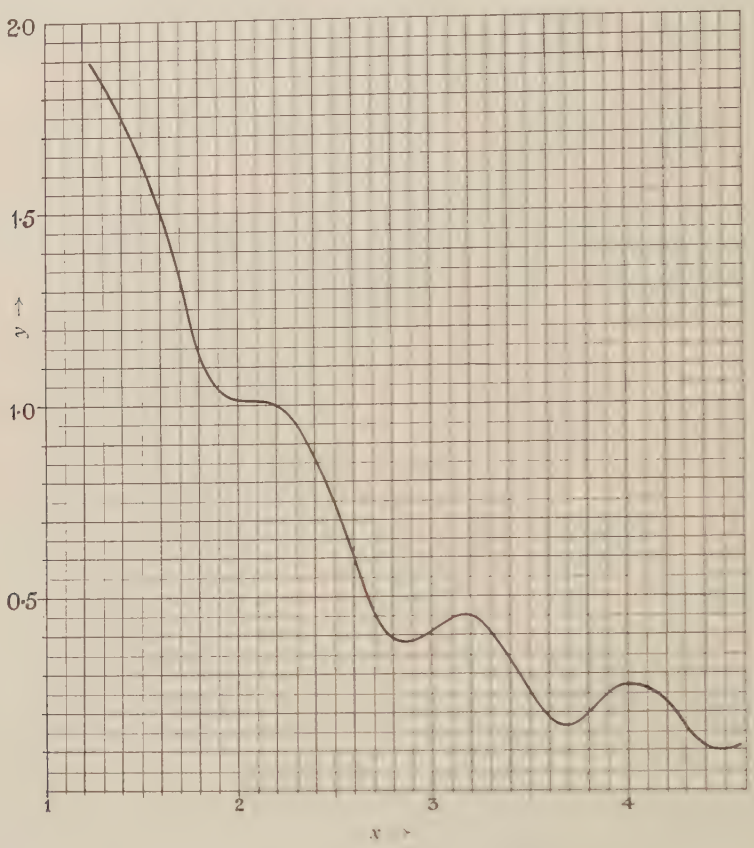


Figure 8.

$$y = \left| \int_{x=2\cdot15}^{x} e^{-i\frac{\pi}{2}u^2} du \right| = f(x),$$

where $x = \sqrt{\frac{3}{2}} \left[1 - \xi' \frac{\cos \psi_0}{A \rho_0 L} \right]$

and θ_0 is angle of incidence;
ψ_0 is angle of diffraction;

L, length of grating;
R_0, distance of source from grating;

$$A = \frac{1}{2} \left\{ \frac{\cos^2 \theta_0}{R_0} + \frac{\cos^2 \psi_0}{\rho_0} \right\};$$

ρ_0, distance of centre of image from grating;
ξ', co-ordinate of field point.

A graph of y as a function of x, plotted from equation (14) is given in figure 8.

It will be seen that the turning value given by equation (15) for $n = 1$ is a maximum, with one minimum very close on the left of it. There is almost an inflection at

$$x = \sqrt{\tfrac{3}{2}} + \sqrt{\tfrac{2}{3}}, \text{ i.e. } n = 1.$$

The values of x for $n=2$, 3, 4, are all definite minima with maxima following them. There is no turning value for a value of x between $\sqrt{\frac{3}{2}}+\sqrt{\frac{2}{3}}$ and $\sqrt{\frac{3}{2}}$, so the central Airy band is of almost exactly the same width as in the ordinary telescope theory, viz. λ/V.

Phase-swinging. When $x=\sqrt{\frac{3}{2}}$ the phase of the Fresnel integral is $-42^\circ 33'$; when $x=\sqrt{\frac{3}{2}}+\sqrt{\frac{2}{3}}$ it is $-22^\circ 49'$, so that in passing across the central band from the centre to the minimum there is a phase-swing of about 20°, corresponding to $\frac{1}{18}$ of a wave-length. The parabola defined by equation (4) is thus not strictly an isophasic. The effect of this is that the true isophasic over the first band is not the parabola but a curve very close to it and differing from it at the first inflection by $\lambda/18$.

Now under the conditions in which these gratings are used at grazing angles, the actual half-breadth of the central image is some 6000 to 7000 wave-lengths. For example, if $\psi_0=87^\circ$, $L=0{\cdot}725$ mm. and $\rho_0=10''$,

$$2V=L\cos\psi_0/\rho_0=0{\cdot}000147,$$

and hence

$$\xi_1=\lambda/0{\cdot}000147=6800\lambda.$$

Consequently the parabola and the true isophasic are very nearly identical, in spite of the phase-swing.

The Poynting vector at any point of a (true) isophasic is orthogonal to it. If then we suppose the parabola to be the true isophasic, the error involved is that of neglecting the cosine of the angle between the normals to the parabola and the true isophasic at the point in question, as a factor. Since the two curves are so close to each other, the error in taking this cosine is obviously quite negligible.

We are therefore justified in squaring the ordinates of the curve in figure 8, and taking these numbers as proportional to the intensity of the field at different values of x.

Resolution. A very simple calculation now shows that if we place the images so that the maximum of one falls on the first minimum of the other, the combined intensity between the two maxima falls to 82 per cent of either maximum. From figure 8 we see that the values of y (the amplitude) are given by:

$$y=1{\cdot}01 \text{ when } x=2{\cdot}04,$$
$$y=1{\cdot}89 \text{ when } x=1{\cdot}225,$$
$$y=1{\cdot}37 \text{ when } x=1{\cdot}63.$$

The combined intensity-maxima are $(1{\cdot}89^2+1^2)$, which $=4{\cdot}57$; and the combined intensity-minima are $2\times 1{\cdot}37^2$, which $=3{\cdot}76=82$ per cent. This value corresponds to the 81 per cent of the Rayleigh rule. Consequently the conditions are almost exactly the same as those defined by the Rayleigh rule, and the breadth of the central band is the same as that of the Airy band; consequently we must have the same, the ordinary, resolution limit, Mr; in fact

$$-b\cos\psi_0.\delta\psi_0=r\delta\lambda_0,$$

i.e.,

$$b\cos\psi_0\cdot\frac{\lambda}{L\cos\psi_0}=r\delta\lambda_0, \text{ by equations (17) and (17}a\text{)},$$

i.e.,

$$\lambda_0/\delta\lambda_0=rM \qquad \text{.....(18),}$$

where M is the total number of lines on the grating.

M

It is rather remarkable and unexpected that the central band formed by such a grating should have the same width as the usual Airy band, and in spite of not being zero at the edge of the band should give the same resolution formula. The explanations are as follows. (1) Both bands are diffraction bands and hence depend on Fresnel integrals, in spite of the fact that the Airy image theory is most easily given as a first-order theory. (2) Since intensity is not zero at the first minimum, the combined maxima are considerably increased. Spectroscopists* do not anticipate so high a resolving power as *Mr*. It must be carefully borne in mind that theory only anticipates it provided the grating-length is chosen to comply exactly with the condition that the range of the Fresnel integral, for the centre of the image, shall be *exactly* $2\sqrt{\frac{3}{2}}$.

Minimal conditions. Just as there is an optimum length so there is a worst length—in fact, it would be just as easy to design the grating-length to give a minimum at the centre and brighter bands on either side, an image that might be mistaken for a double line. The conditions for this are obvious from figure 6. The choice of the proper length for optimum conditions depends on equation (16*a*), and to secure this either the grating must be ruled with the proper number of lines or the two slits must be chosen so as to pass the required very narrow pencil of light only.

Dispersion.
$$\frac{d\psi}{d\lambda} = \mp \frac{r}{b \cos \psi_0} \qquad \text{.....(19),}$$

the upper or lower sign being taken according as ψ_0 is less than or greater than θ_0. The dispersion is unusually high when ψ_0 is nearly 90°, as it is in actual cases, but it still falls far short of the dispersion of a crystal grating.

Numerical examples. The figures for three cases are given in table 1.

The second column in table 1 is calculated for a grating of moderate pitch (1/15000″), an angle of incidence θ_0 equal to 89°, and a wave-length λ equal to 20 A. The grating is very short, only 0·725 mm. in length, and the number of lines is 427. The third column shows what happens if this grating is used at $\lambda = 40$ A., in which case of course the grating is not of optimum length. The central intensity is almost the same; the width of the central band increases by 43 per cent but the resolution is the same as before. The dispersion falls off by 27 per cent. The structure of the band does not alter radically over the range $\lambda = 20$ to 40 A. Consequently a grating designed for optimum conditions at $\lambda = 30$ A. might be expected to work satisfactorily between $\lambda = 10$ A. and $\lambda = 50$ A. The fourth column is calculated for a very fine pitch ($p = 1$ 45000″). Comparing this column with the second, we see the glancing angle of reflection is nearly doubled; the angular aperture of the pencil is much as before; the length of the grating is 0·445 mm., 61 per cent of its former value; the number of lines goes up by 84 per cent; the central band hardly changes. But the dispersion goes up by 90 per cent and the resolution by about the same amount.

Overlap of images. The intensity-bands we have discussed arise from a Huyghens

* For example, Siegbahn[3], p. 2.

Table 1

Pitch, b	16,950A. or 1/15000″	16,950A. or 1/15000″	5650A. or 1/45000″	Data
Order, r	1	1	1	—
Dimensions, R_0 and ρ_0 (in.)	10	10	10	—
Grazing-incidence angle, γ_0 (deg.)	1	1	1	—
Wave-length, λ (A.)	20	40	20	—
Grazing-reflection angle, δ_0	2°57′	4°4′	4°56′	By Eqn. (8*a*)
Angular aperture, u (sec. of arc)	30	40	31	,, (17*a*)
Length, L (mm.)	0·725	0·725	0·445	,, (3) and (16*a*)
Number of lines, M	427	427	785	—
Width, $2\xi_1'$, of central band (A.)	$27{\cdot}3 \times 10^4$	$39{\cdot}2 \times 10^4$	$26{\cdot}5 \times 10^4$	,, (17)
Resolution, $\lambda/\delta\lambda$	427	427	785	,, (18)
Dispersion per angstrom, $d\psi/d\lambda$	$1{\cdot}145 \times 10^{-3}$	$0{\cdot}831 \times 10^{-3}$	$2{\cdot}06 \times 10^{-3}$	,, (19)
Range of Fresnel integral	$\pm\sqrt{\frac{3}{2}}$ or $\pm 1{\cdot}225$	$\pm 1{\cdot}16$	$\pm 1{\cdot}225$	—
Relative intensity	1	0·99	1	—

source on the extended source of radiation. The slits must be chosen so that the exposed line of the radiant source is sufficiently narrow to prevent serious displacement of the images formed by the elementary Huyghens sources of which the source line consists. It would be easy to destroy the whole advantage of this type of band by permitting too wide a line on the source to be exposed simultaneously. The exact conditions required will be evident from the calculations already given. The extreme Huyghens sources should be near enough to ensure that their respective images shall not be relatively displaced on the film by more than say 5 per cent of the central band-width. It may be impossible to secure this; but anything more will clearly lead to the production of a more diffuse band on the film.

Validity of the fundamental grating formula. It has been suggested that wave-lengths determined from the formula

$$b\,(\sin\theta_0 - \sin\psi_0) = \pm r\lambda$$

by plane-grating methods in the soft X-ray region are not quite correct. We have found no evidence for this view. We have shown that the centre of the image lies

exactly in the direction given by this rule and it is the angular position of the centre of the image that it is the object of experiment to determine.

Third-order effects. Our theory includes only second-order terms in the expansion of $(R+\rho)$ but we can readily show that, in the conditions contemplated in table 1, the third-order term is negligible. The third-order coefficient is a_3, where

a_3

$$a_3 = \frac{1}{2}\left[\frac{\sin\psi\cos^2\psi}{\rho_0^2} - \frac{\sin\theta\cos^2\theta}{R_0^2}\right] \qquad \text{.....(20).}$$

The maximum length that this term contributes to $(R+\rho)$ is a_3x^3, where x is the half-length of the grating or, in terms of the wave-length, $(a_3x^3/\lambda)\lambda$.

γ Now $b(\sin\theta - \sin\psi) = r\lambda$; hence, if $\theta = 90^\circ - \gamma$

δ and $\psi = 90^\circ - \delta$,

$$\tfrac{1}{2}b(\delta^2 - \gamma^2) \approx r\lambda.$$

If $R_0 \approx \rho_0$

$$a_3 \approx \frac{1}{2\rho_0^2}[\delta^2 - \gamma^2]$$

$$\approx r\lambda\, b\rho_0^2$$

and

$$a_3x^3/\lambda \approx \frac{r}{b\rho_0^2}x^3$$

$$\approx \frac{r}{b\rho_0^2}\cdot\frac{L^3}{8}$$

$$\approx \frac{r}{8}\left(\frac{L}{\rho_0}\right)^2 L$$

$$\approx \frac{1}{8}\left(\frac{L}{\rho_0}\right)^2 Mr,$$

where M is the total number of lines. This formula is true whether the grating is of optimum length or not, provided that $R_0 \approx \rho_0$.

If $\rho_0 = 10''$, $L < 1$ mm. and $Mr < 800$, then

$$\frac{a_3x^3}{\lambda} < \frac{1}{645},$$

so that the *maximum* error in omitting this term is only some 645th part of a wave-length. It is difficult to see how the inclusion of this term could materially affect our results.

Note added July 20, 1935

In the July (1935) number of the *Proceedings* of this Society, Petrie[5] gives very interesting details of a single-slit plane spectrograph. In this apparatus a very obliquely inclined X-ray target is used as the source slit, and the width of the beam of light leaving the grating is limited by an adjustable slit placed very close to the end of the grating and at right angles to the grating-face. The grating itself is not ruled for optimum conditions—if it were, the beam leaving the grating would be already focused so that the slit would be superfluous. The image photographed with this instrument is the diffraction image of the slit used.

REFERENCES

(1) R. W. Wood and T. Lyman. "Improved gratings for vacuum spectrographs." *Phil. Mag.* p. 310 (July, 1926).
(2) H. M. O'Bryan. "Reflecting power and grating efficiency in the extreme ultra violet." *Phys. Rev.* **38**, 2nd series, 1, p. 32 (July, 1931).
(3) M. Siegbahn and T. Magnusson. "Zur Spektroskopie der ultraweichen Röntgenstrahlen, I." *Z. Phys.* **62**, Hefte 7 and 8 (1930).
(4) J. A. Prins. "Beiträge zur Plangitterspektroskopie der ultraweichen Röntgenstrahlen." *Z. Phys.* **69**, Hefte 9 and 10 (1931).
(5) D. P. R. Petrie. "Comparison of X-ray wave-lengths by the plane-grating vacuum spectrograph, and the structure of the *K* line of carbon." *Proc. phys. Soc.* **47**, 626 (1935).

621.318.423:621.317.334

A DYNAMOMETER NULL METHOD OF MEASURING THE INDUCTANCE AND THE EFFECTIVE RESISTANCE OF IRON-CORED CHOKES CARRYING DIRECT CURRENT

By GEOFFREY D. PEGLER, M.Sc.

Received March 8, 1935. *Read June* 21, 1935

ABSTRACT. An alternating potential-difference is applied to the terminals of two parallel circuits, one of which contains the fixed coils of a reflecting electrodynamometer, the inductance to be measured and a known variable resistance, while the other contains the moving coil of the dynamometer in series with a condenser of capacity K and a second known variable resistance. The former resistance is fixed during a determination, while the latter is varied until the phase-difference between the currents in the two branches is exactly $\frac{1}{2}\pi$, in which case there is no deflection of the moving coil. The inductance is then given by $L=KR_KR_L$, where R_K and R_L are the total effective resistances of the capacity and inductance branches respectively. A second balance is now obtained with the value of the first resistance altered by a suitable and known amount, the second resistance being altered accordingly. From the data of the two balances both the effective resistance (which is included in R_L) and the inductance may be calculated. The method is equally applicable when direct current is allowed to flow through the inductance. It is accurate over a range of inductances from a small fraction of a henry to that of the largest chokes. Examples of measurements are given.

§ 1. INTRODUCTION

In the discussion which followed a paper published in these *Proceedings*[1], a null or no-deflection method of using an electrodynamometer for measuring the inductance of iron-cored chokes was suggested. The present paper carries out this suggestion, and provides thereby an extremely simple means of accurately determining the inductance of a choke in terms of resistance and capacity, and also its simultaneous a.-c. resistance with or without the superposition of direct current.

The change of zero due to imperfect elasticity of the suspension, inherent in a deflectional method, is absent, and the absence of a deflection also dispenses with errors introduced by mutual inductance between the fixed and moving coils of the electrodynamometer. The balances necessary for a determination may be approached by the adjustment of a single variable. Since the frequency is not involved in the fundamental equation, the method is independent of the wave-form of the alternating current used, in so far as the quantities measured are themselves independent of frequency.

§ 2. THEORY

If alternating currents of the same frequency are flowing in both the fixed and moving coils of an electrodynamometer there will be a deflection of the moving coil unless the phase-difference between the two currents is exactly 90°, the sign of the deflection changing as the phase-difference passes through that value. The absence of a deflection signifies that the two currents are in exact quadrature and this fact may be applied to the measurement of inductance in the following way.

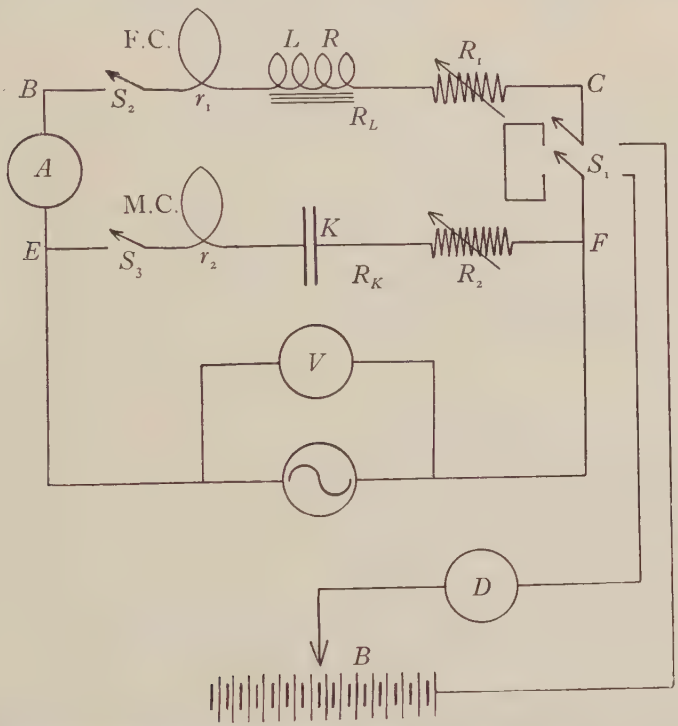


Figure 1 *a*. Diagram of circuits. *F.C.* is the fixed coil and *M.C.* the moving coil of the electrodynamometer; *LR* is a choke, *K* is a condenser, and R_1 and R_2 are variable resistance boxes. R_L and R_K denote the total resistances of the branches *BC* and *EF*. *A* and *D* are a.-c. and d.-c. milliammeters respectively, *V* is an a.-c. voltmeter; S_1, S_2, S_3, are switches, and *B* is an h.-t. battery.

An alternating voltage is applied, figure 1 *a*, to two circuits in parallel, one containing the fixed coils of the dynamometer in series with the choke of inductance L, to be measured, and a known variable resistance; the other circuit including the moving coil, in series with a condenser of capacity K and a second variable and known resistance. The current will lead the applied volts in the capacity branch by the phase angle ϕ', and will lag in the inductance branch by the angle ϕ''. Assuming simple harmonic currents and voltages, and denoting the total resistances in the inductance and capacity branches respectively by R_L and R_K and the r.m.s. values of the currents by I_L and I_K, we shall have figure 1 *b* as the vector diagram representing the phases of the currents and the applied voltage.

L
K
ϕ', ϕ''
R_L, R_K
I_L, I_K

When, by the variation of R_L or R_K, quadrature between I_L and I_K is obtained, we have

$$\tan \phi' \tan \phi'' = 1.$$

Since $\phi' = \tan^{-1} 1/pKR_K$, and $\phi'' = \tan^{-1} pL_L R_L$, the condition for no deflection of the moving coil is therefore

$$pL/R_L \times 1/pKR_K = 1$$

or

$$L = KR_L R_K \qquad \ldots\ldots(1),$$

a simple relation requiring no measurements of frequency, current, or applied voltage.

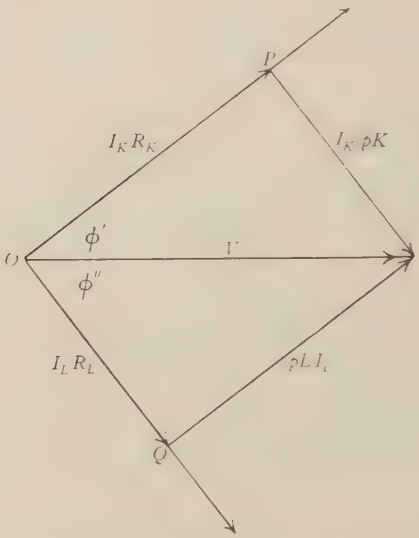


Figure 1*b*. Vector diagram representing the relative phases of the currents through the branches *BC* and *EF* of figure 1*a*. *OP* represents the phase of the condenser current leading the applied voltage V by ϕ', *OQ* the phase of the choke current lagging by ϕ'', while $\phi' + \phi'' = 90°$.

δ, k If, when the balance is being obtained, the phase defect from perfect quadrature is δ, the torque on the moving coil may be written $kI_L I_K \sin \delta$, where k is an instrumental constant. The differential coefficient of sin δ, has its maximum value when δ is zero, with the desirable result that the instrument will be most sensitive R_1, R_2 to phase-differences at the null point. If R_1 and R_2 denote the values of the re- r_1, r_2 spective box resistances in the inductance and capacity branches, r_1 and r_2 the R resistances of the fixed and moving coils of the electrodynamometer, R the a.-c. R_0 resistance of the choke, and if further we denote the sum of R_1 and r_1 by R_0, we have

$$R_L = R_1 + r_1 + R = R_0 + R \text{ and } R_K = R_2 + r_2.$$

The formula in greater detail thus becomes:

$$L = KR_K (R_0 + R) \qquad \ldots\ldots(2).$$

To find R, the value of R_0 is changed by altering R_1, and R_K is correspondingly R_K', R_0' altered until a new balance is obtained. Denoting the new values by R_K' and R_0' we have

$$L = KR_K' (R_0' + R) \qquad \ldots\ldots(2'),$$

and the effective resistance is given by

$$R=(R_K R_0-R_K' R_0')/(R_K'-R_K) \qquad \ldots\ldots(3).$$

R having been determined from equation (3), the inductance L may now be calculated by substituting this value in equation (2). The value of L so found includes the inductance of the fixed dynamometer coils, which is quite negligible when iron-cored chokes are being measured.

If the alternating voltage is supplied through a transformer with a secondary of low d.-c. resistance, direct current may be injected into the branch containing the choke without affecting the moving coil.

§ 3. EXPERIMENTAL CIRCUIT

To investigate the method experimentally, a reflecting electrodynamometer was constructed having two vertical parallel fixed coils of 500 turns each, of mean diameter 2·5 inches, and about 2·5 inches apart; and suspended centrally between these a moving coil of 300 turns, of mean diameter 1 inch, carrying a small mirror of about 1 metre radius. The number of turns in the coils is only limited by the heating effect, which must not be permitted to generate convection currents sufficient to disturb the zero of the suspended system. With the coils specified this effect was undetectable, while the accuracy of measurement secured was ample for the purpose, see table 1.

The suspension was a phosphor-bronze strip of length 3 in. and cross-section 0·007 × 0·0005 in., and a similar nearly-taut strip led the current out to the lower contact. The period of swing was about 6 sec., and a damping vane of mica attached below the coil rendered the movement nearly deadbeat. The resistance of the moving coil including the strip was 35 Ω., and that of the two fixed coils in series was 60 Ω.

K was taken from a calibrated, decade, 1-μF. condenser subdivided down to 0·001 μF., and R_1 and R_2 were taken from non-inductive dial resistance boxes giving 0 to 10,000 Ω.

The alternating voltage V was taken from the secondary of a transformer of low d.-c. resistance, the supply to the primary being regulated by a potential-divider, preferably non-inductive, connected across 220-volt 50-cycle mains.

By the aid of the double-pole throw-over switch S_1, figure 1*a*, direct current can be injected into the choke circuit from a battery of small accumulators, this current being read on the d.-c. milliammeter D. The thermal milliammeter A registers the r.-m.-s. value of the total current. Alternatively it is possible, by means of a thermionic voltmeter, to read the alternating current component separately. Both the inductance and capacity circuits can be broken independently by means of the switches S_2 and S_3.

§ 4. EXPERIMENTAL PROCEDURE

The electrodynamometer is set up with the plane of its moving coil at right angles, as judged by eye, to that of the fixed coils. The terminals of the moving coil are short-circuited, while alternating current is passed through the fixed coils.

The suspension is adjusted until there is no change of zero of the moving coil when the current is switched on or off. The mutual inductance between the coils is now zero.

To test a choke with alternating current alone, the switch S_1 is closed to the left, V is set at a low value, S_2 is closed and S_3 left open to prevent needless deflection of the moving coil. V is increased until the alternating current at which the test is to be made is registered by the milliammeter A. S_3 is then closed momentarily and R_2 is adjusted, and this process is repeated until the position of the spot on the scale remains unaffected by the opening or closing of S_3. After the values of the box resistances R_1 and R_2 have been noted, R_1 is increased by a suitable amount and the foregoing procedure is repeated. From the data of these two successive balances both the inductance and the effective resistance of the choke may be calculated. Of course, the increase in R_1 should be of the same order as the effective resistance of the choke as found by a preliminary determination.

The switching necessary for a determination with direct current superposed is evident from the circuit diagram. The required direct current, registered by the d.-c. milliammeter D, is obtained by adjusting the wanderplug connexion to the battery, a small completing adjustment being made by varying R_1. V is altered until the r.-m.-s. sum of the required d.-c. and a.-c. currents is registered by A, S_3 is closed, and the procedure for two balances is followed as before, save that the value of the direct current must be kept unaltered by a readjustment of the battery connexion when R_1 is altered for the second balance. Preferably the a.-c. component is measured separately with a grid-leak type of thermionic voltmeter connected across a known resistance in the choke circuit.

If the currents through the choke are kept constant during a series of tests, and the values of R_0 and $1/KR_K$ for each balance point are plotted as ordinates and abscissae respectively, a straight-line graph intersecting the Y axis at $-R$ will result. The slope of the graph gives the value of L, so that the values of both L and R may quickly be obtained. They may also be calculated, of course, from the data by applying equations (3) and (2).

§ 5. EXPERIMENTAL RESULTS

Figures 2*a* and 2*b* represent a series of graphs plotted as described above for two chokes, and indicate the degree of dependence of their inductance and effective resistance on the value of the superposed direct current. The graphs in figure 2*a* refer to a choke of good design, rated at 20–30 henries, with a direct current-carrying capacity of 50 mA. and of which the d.-c. resistance was 387 Ω. The graphs in figure 2*b* appertain to a choke of inferior quality rated by its manufacturers at 20 henries, the d.-c. resistance in this case being 340 Ω. From figure 2*a* it will be observed that the effective resistance, as shown by the intercept on the vertical axis, varies slightly with the superposed direct current in the same way as the inductance, showing first an increase followed by a decrease. Both the inductance and the effective resistance of the second choke when no direct current is passing are seen to have very high values (117 H., 16,700 Ω.), which however fall rapidly

when direct current is superposed. Thus when 12 mA. is flowing the values become 22·19 H. and 1400 Ω. The very large drop in the effective resistance shows

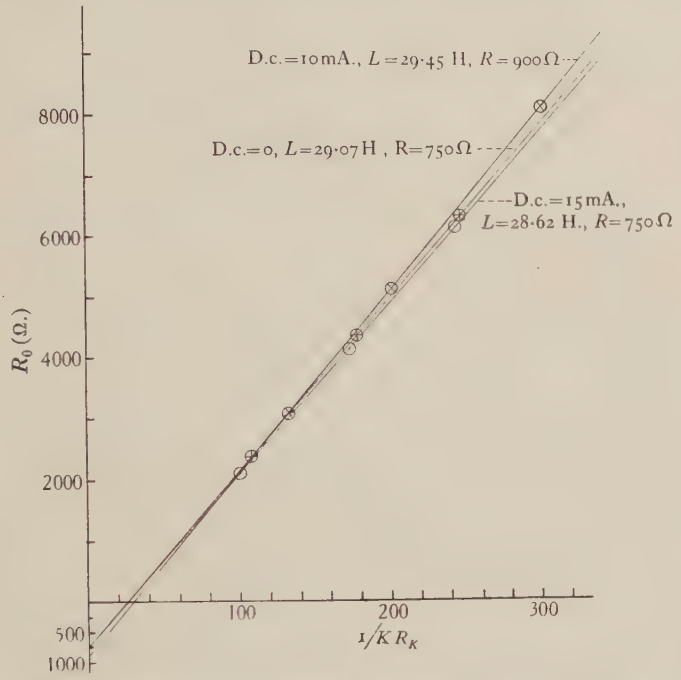


Figure 2*a*. Results for a choke of good design. Frequency of supply 50 c./sec. K was constant and equal to 1 μF. The a.-c. was in all cases 3 mA.

how considerable may be the effect of a polarizing current on hysteresis and eddy current losses, to which the resistance is mainly due. A typical set of results, from which one of the graphs in figure 2*b* is drawn, is given in table 1.

Table 1. $K=1\,\mu$F., a.-c. = 3 mA., d.-c. = 8 mA., frequency = 50 c./sec., $r_1=60\,\Omega.$, $r_2=35\,\Omega.$

V (V.)	R_1 (Ω.)	R_2 (Ω.)	R_0 (Ω.)	R_K (Ω.)	$1/KR_K$
31	1190	9570 ± 10	1250	9605	103·1
34	4280	4842 ± 3	4340	4877	204·9
39	6600	3550 ± 3	6660	3585	278·3
42	8150	3023 ± 2	8210	3058	322·0

Hence, from the graph, $R=1900\,\Omega.$ and $L=30{\cdot}68$ H. when the superposed d.-c. is 8 mA. From column 3, showing the accuracy of setting, it is seen that the individual value of L can be regarded as correct to about 1 part in 1000.

To illustrate the applicability of the method to the measurement of small inductances, the choke was removed from the circuit and the inductance of the fixed coils of the electrodynamometer alone measured. In this case R_1 was kept constant, balances being made by giving K different values and varying R_2 until the deflection of the moving coil in each case was zero.

The graph connecting values of K with the corresponding values of $1/R_L R_K$ proved to be very closely linear and to pass through the origin. The value of L deduced was 0·049 H.±1 per cent. A determination by a bridge method gave 0·0505 H.

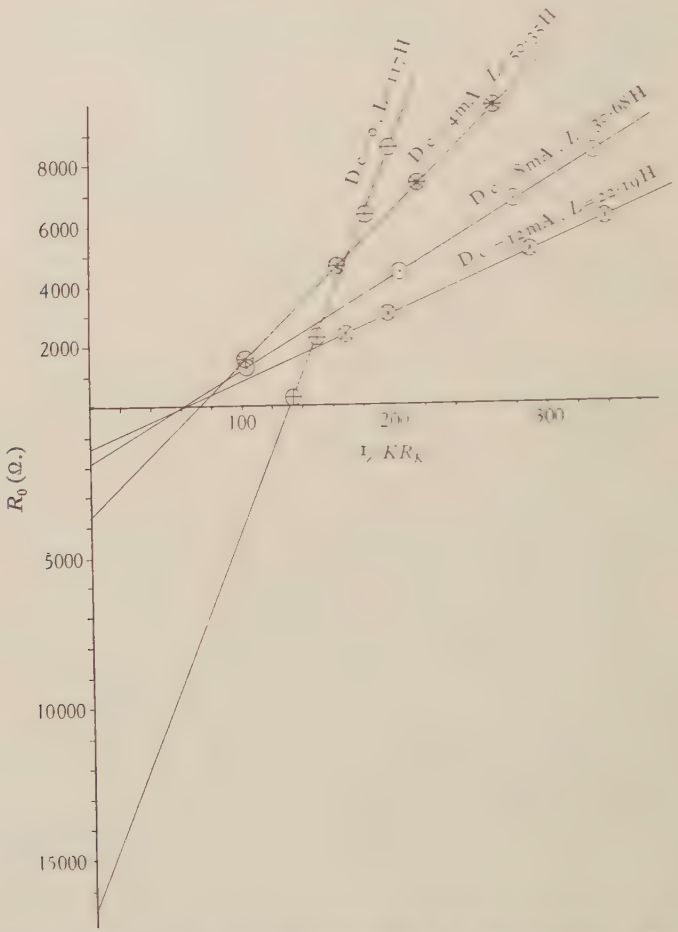


Figure 2*b*. Results illustrating the properties of a choke of inferior design. Frequency of supply 50 c./sec. The slope of the straight lines is a measure of the inductance; the intercept below the origin on the axis of ordinates is the effective resistance of the choke.

§ 6. RANGE AND ACCURACY

The method serves for a wide range of values of inductance. For chokes having very high impedance the voltage V must, of course, be correspondingly increased in order to provide the alternating current required for the test. To keep the current in the moving coil within suitable limits, the impedance of this branch must likewise be increased, and this may be effected by decreasing K and, at the same time, increasing R_K. Conversely, when small inductances are being measured V must be small, and to maintain high sensibility the impedance of the moving coil branch must be correspondingly decreased by increasing K and decreasing R_K.

Errors may arise from two sources: (i) the inductance of the moving coil in the capacity branch, and (ii) the self-capacitance of the choke itself. The importance of both effects will depend on the frequency of the alternating current used, and on 50-cycle mains both effects will be negligibly small. In regard to (i) the inductance of the moving coil was 0·0032 H. and the largest value of K used was 1 μF. The exact formula for balance is

$$p\,(L+l_f)/R_L \times (1/pK - pl_m)/R_K = 1,$$

where l_f and l_m are the respective inductances of the fixed and moving coils of the electrodynamometer. l_f, l_m

Thus
$$(L+l_f) = KR_KR_L/(1-p^2l_mK) \qquad \ldots\ldots(4),$$

so that, at a frequency of 50 c./sec., the error in the measurement of L due to neglecting the term l_m is less than 1 in 3000. For smaller values of K the error is correspondingly less and in all cases at this frequency it may be neglected.

In making a measurement at the higher audio frequencies a knowledge of the frequency is required and the correction, as shown in equation (4), must be applied when the inductance is calculated. The linearity of the graphs drawn with co-ordinates $1/KR_K$ and R_0 is unaffected by frequency, and the effective resistance of the choke under test may be found from the graph or by equation (3) without correction. The choke for which the graphs in figure 2*a* were drawn, at a frequency of 50 c./sec., was tested at frequencies of 768 and 1000 c./sec., a valve oscillator being used as a source. The effective resistance was found to increase considerably, becoming 11,000 Ω. at 768 c./sec., and 18,000 Ω. at 1000 c./sec.

The large value of effective resistance of some chokes in comparison with their d.-c. resistance has been already remarked. It is clear that the frequently-employed method of calculating the impedance of a choke from its inductance and d.-c. resistance is unjustifiable.

In regard to (ii), if the self-capacitance of an inductance be regarded as a small condenser C in parallel with it, the effective value L' of the inductance will be given by $L' \doteqdot L/(1-p^2CL)$ and the effective resistance R' by $R' \doteqdot R/(1-2p^2CL)$, where L and R are its values at very low frequency.

In the measurement of inductance we are not, however, directly concerned with the self-capacity term, since it is the actual effective inductance and effective resistance under the known applied frequency that are required.

§ 7. ACKNOWLEDGMENTS

The above work was carried out in the Physics Research Laboratory of the Sir John Cass Institute, London. My best thanks are due to Dr D. Owen, upon whose suggestion the method is based, for his advice and interest throughout the work.

REFERENCE

(1) Willoughby. *Proc. phys. Soc.* **46**, 292 (1934).

DISCUSSION

Dr D. Owen. I believe the method will prove its value, serving as a quick and accurate means of measuring not only the self-inductance of large inductances, but also their effective resistance to alternating currents. The dynamometer is an instrument whose merits deserve the attention of engineers interested in accurate measurements, providing as it does an indicator applicable over the full range of audible frequencies and even beyond it. Its construction, suitable data of which for the measurements under present consideration are given by the author, offers no serious difficulty, nor is any special care needed in its adjustment for use.

It may be noted that, as the formula indicates, a null reading may be attained not only by varying the resistances, but also by varying the capacity. Now that accurate decade mica condensers are available at prices comparable with those of resistance boxes this consideration might be applied also to many bridge methods for the measurements of inductance or capacity. For instance, the use of a variable capacity eliminates the usual objection to Maxwell's inductance-capacity bridge, and should result in a return to that method, so is desirable otherwise on account of its fundamental simplicity.

Dr L. Hartshorn. In comparing methods for the testing of choke coils carrying both alternating and direct current, perhaps the most important consideration in practice is the possibility of making the test with the choke working at full load, i.e. carrying the largest possible direct current. It was this consideration which led me to choose Hay's bridge for such measurements some years ago, since this bridge satisfies the condition to a remarkable degree. The measurements mentioned in the paper are limited to chokes carrying very light loads, and some modification of procedure would probably be required for measurements under actual working conditions. It would be of interest to hear what are the capabilities of the method in this respect.

Mr A. Campbell. If the current in the condenser branch contains no harmonic components, then any such components as may occur in the L branch have no effect on the wattmeter reading; in this case the equation $L = KR_L R_K$ is quite valid, and the results obtained should be the same as if a method with a tuned detector were used. But if the supply voltage is not purely sinusoidal, in general both the branch currents will contain harmonics; and although a setting to give zero deflection can always be made, it seems impossible that the same equation should still hold. The author might find it interesting to explore such a case by testing ironless coils with a supply voltage of much-distorted wave form.

Author's reply. The only modification of the present method for the testing of chokes under full load conditions is that which was employed by Dr Hartshorn himself, namely the inclusion in the choke circuit of a fixed resistance specially wound to dissipate the heat generated at large values of direct current. In this case the two necessary balances would be made first with no other resistance than tha

of the choke itself in the inductance branch, and secondly with the fixed resistance included.

In the dynamometer method there is only one condition of balance to be satisfied, and this gives a distinct advantage over a bridge method with its two necessary balance conditions which are often not independent. The Hay bridge suffers from the latter disadvantage and also from the fact that its balance equations include the frequency, a fact of fundamental importance when the measurement of a choke with a core of variable permeability is to be made, since in this case the choke current will contain a high percentage of harmonics. Another advantage of the present method is that, when a balance point is being approached, the current need only be switched on for long enough to observe the direction of deflection of the moving coil, which indicates the sense in which the variable is to be altered. By this means the heating-effect in the resistance is reduced to a minimum.

I am obliged to Mr A. Campbell for pointing out a source of possible error due to harmonics in the supply voltage. It was shown in Mr Willoughby's paper, in which the deflectional method of using the electrodynamometer for such measurements was introduced, that to a first approximation the presence of harmonics cancelled out. In the null method, however, harmonics in the capacity branch will certainly tend to affect the accuracy of the observed readings, the amount depending on the permeability of the core, the value of the alternating current used, and the percentage of harmonic content in the supply voltage. I am indebted to Mr Campbell for his suggestion.

537.226

THE BREAKDOWN OF DIELECTRICS UNDER HIGH VOLTAGE, WITH PARTICULAR REFERENCE TO THERMAL INSTABILITY[1]

By S. WHITEHEAD and W. NETHERCOT

Received January 18, 1935. *Read in title June* 7, 1935

ABSTRACT. Provided a sufficiently wide range of external conditions can be applied, dielectrics exhibit both a thermal and non-thermal type of breakdown. The electric strength for thermal breakdown can be calculated from the stable electrical properties of the dielectric and a simplified theory is given in the paper which is shown to agree with experiment for a number of dielectrics for different conditions of temperature, time of application of stress, etc. A theory of ionization coefficients is developed to explain conductivity phenomena with a.-c. and d.-c. and ionization potentials are deduced. It is, however, shown that these phenomena do not appear to bear a direct relation to the possible types of non-thermal breakdown and that the ionization potentials have not the same physical significance as the corresponding quantities in gases. Experimental results relating to the type of non-thermal breakdown which occurred in the present investigations are given.

§ 1. INTRODUCTION

Origin of the investigation. The theory of the breakdown of dielectrics through thermal instability has already been studied fairly extensively[2], but mainly by means of breakdown tests in which only the field-strength, at which a marked and usually sudden instability occurs, has been noted and its variation with temperature, for example, has been compared with the corresponding variation of the conductivity. The absolute values of electric strength have often not agreed very closely with theory, while the materials examined and the ranges of temperature employed have been of a rather special character. In the work here described the behaviour of a number of well-known dielectrics was studied both before and during breakdown, at temperatures at which they are normally sufficiently stable to be employed for insulation purposes.

The main objects of the investigation were to determine whether the processes occurring before breakdown are compatible with a simple theory of thermal instability and whether the electric strength can be calculated from a knowledge of these processes in the stable range, even though other processes may occur at the instant of breakdown. In addition, it was desired to study the limiting features which determine the range of application of the theory with respect to temperature, time of stress and other factors.

General thermal instability: simplified theory. The energy-losses which occur in a dielectric under electric stress cause a temperature-rise distributed in a certain way. If the energy-losses increase with temperature and electric stress, a condition

may occur when the generation of heat rises more rapidly with temperature than the thermal dissipation, and the temperature at some point or region will rise more and more rapidly, eventually causing a loss of insulating properties or an electric discharge. One of the simplest cases is that of a sheet of dielectric of thickness d in a uniform field with plane electrodes which cool the dielectric surface at a constant rate of λ watts per degree C. per cm.^{2} Suppose also that the energy-losses W at $\theta°$ C. above the ambient temperature are given by

$$W = E^2 p_0 e^{\gamma\theta} \text{ watts/cm.}^3 \quad \text{.....(1),}$$

where p_0 is the specific energy-loss (a.-c. conductivity) at the ambient temperature, E the electric field and γ a constant. Utilizing Fock's[3] analysis, we find that the maximum voltage V which can be applied to the dielectric without causing thermal instability is as follows:

$$V = \sqrt{(2/p_0\gamma\rho)} \sin\alpha \exp\left[-\alpha\cos^2\alpha/(\alpha+\sin\alpha\cos\alpha)\right] \quad \text{.....(2),}$$

where ρ is the thermal resistivity of the dielectric in degrees C. per watt and α is given by

$$\lambda d\rho/2 = \sin\alpha \sec^3\alpha\,(\alpha+\sin\alpha\cos\alpha) \quad \text{.....(3).}$$

We observe that $\lambda d\rho/2$ is P_i/P_e, where P_i is half the thermal resistance of the dielectric and P_e is the thermal resistance of an electrode. In other words, α is a function of P_i/P_e or the ratio of the internal to the external thermal resistance of each symmetrical half of the arrangement. We also observe that if the internal thermal resistance is negligible, then the maximum voltage V_m for stability is given by

$$V_m = \sqrt{(\lambda d/2\, p_0\gamma e)} \quad \text{.....(4),}$$

whence

$$\frac{V}{V_m} = \sqrt{\left[\frac{2\exp\left(\frac{1}{2}\sin 2\alpha - \alpha\cos 2\alpha\right)}{\sec^2\alpha + (\alpha/\sin\alpha)\sec^3\alpha}\right]} \quad \text{.....(5).}$$

Equation (5) shows that V/V_m, or the effect of the temperature-gradient in the dielectric, is a function of P_i/P_e. We may distinguish two extreme cases. When P_i/P_e is great the critical voltage V finally depends only on the electrical and thermal properties of the dielectric and is independent of its thickness, so that a certain maximum voltage is associated with a given dielectric for $\alpha \to \pi/2$ and $V \to \sqrt{(p_0\gamma\rho/2)}$. The calculation of this limiting voltage is usually fairly easy, but is difficult to verify experimentally owing to the high voltages involved and the intervention of extraneous effects, while in practice it is important only with very high-voltage equipment. Accordingly, if we consider the other extreme when P_i/P_e is small we observe from figure 1 that if $P_i/P_e < 0{\cdot}1$ the error which arises from neglecting the temperature-gradient in the dielectric will not exceed 1·5 per cent, while if $P_i = P_e$ the error is about 12 per cent. Although equation (5) is strictly true only when $W \propto E^2$ or the conductivity is independent of field-strength, yet the values when $P_e > P_i$ do not vary greatly if forms differing from equation (1) and of the more usual type occur. Neglecting P_i we then find the following simple equations for the limit of stability:

$$W = f(E_m, \theta), \quad E_m = V_m/d \quad \text{.....(6),}$$

$$\partial W/\partial\theta = \lambda, \quad \text{.....(7),}$$

θ being the temperature of the dielectric.

In the present paper, in addition to equation (1) which gives the solution (4), two general forms are considered:

$$\sigma = W/E_m^2 = p_0 + aE_m \exp\{(\gamma_0 + \gamma E_m)(\theta - \theta_0)\} \quad \text{.....(8.1),}$$

and

$$\sigma = W/E_m^2 = p_0 + (a_0 + aE_m) \exp\{\gamma_0 (\theta - \theta_0)\} \quad \text{.....(8.2).}$$

These lead to

$$\theta_a = \theta_1 \log_e (\theta_1/e\theta_E) + \theta_0 - \theta_p \quad \text{.....(9),}$$

$\theta_1, \theta_0, \theta_1$ θ_E, θ_p

where θ_a is the ambient temperature, θ_0 a constant, $\theta_1 = 1/(\gamma_0 + \gamma E_m)$ for (8.1) or $1/\gamma_0$ for (8.2), $\theta_E = daE_m^3/\lambda$ for (8.1) or $(a_0 + aE_m) E_m^2 d\lambda$ for (8.2), $\theta_p = E_m^2 dp_0 \lambda$. Generally a sufficient approximation is obtained by writing

$$E_m^2 \sigma = W = p_0 E_m^3 e^{\gamma(\theta - \theta_0)} \quad \text{.....(10),}$$

whence

$$E_m = \{\lambda/ed\gamma p_0\}^{\frac{1}{3}} \quad \text{.....(11).}$$

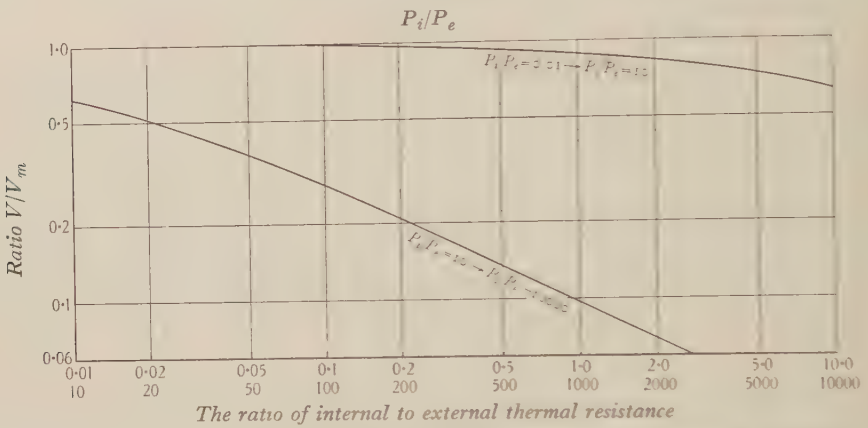


Figure 1. The effect of thermal resistance in the specimen on the thermal instability.

V_m or E_m corresponds to the limit of thermal stability and therefore to the electric strength corresponding to a theoretically infinite time of application, since a slight disturbance would only cause the temperature to rise at an infinitesimal rate. If a greater electric stress is applied, the temperature will rise more rapidly and a critical temperature will be reached in a limited time. We might also consider that there is an intrinsic electric strength which decreases with temperature so that the temperature will rise until this electric strength is equal to the applied field. However, it is usually more in accordance with the facts to consider a limiting temperature θ_c at which, say, some chemical or physical change occurs. The time t for breakdown corresponding to a stress E greater than E_m is given by

θ_c, t

$$\frac{\lambda t}{ds} = \sum_0^\infty \beta^{n+1} \left[\frac{n!}{(n+1)^{n+1}} - e^{-(n+1)\phi} \left\{ \frac{\phi^n}{n+1} + \frac{n\phi^{n-1}}{(n+1)^2} + \ldots + \frac{n!}{(n+1)^{n+1}} \right\} \right] \quad \text{...(12),}$$

β, ϕ, W_{am} W_a, θ_a

where s is the specific heat of the dielectric, $\beta = eW_{am}/W_a$, $\phi = \gamma\theta_c$, and W_{am} and W_a are the energy-losses at θ_a for field-strengths of E_m and E respectively. For equation (1), $\beta = e(E_m/E)^2$ and for equation (10) $\beta = e(E_m/E)^3$, so that $\lambda t/ds = f(E_m^2/E^2)$ and $f(E_m^3/E^3)$ in the two cases respectively. Thus the ratio of the electric strengths for two given times of breakdown is independent of the temperature at which the ratio is taken.

In actual fact, equation (12) may usually be simplified to

$$\lambda t/ds = \sum_{1}^{\infty} (n-1)\,!\,(\beta/n)^n - \beta e^{-\phi} \qquad \ldots\ldots(13).$$

If the ambient temperature is low compared with θ_c, a further simplification is possible, thus

$$\lambda t/ds = \beta\,(1-e^{-\phi})$$

or

$$(E/E_m)^n = eds\,(1-e^{-\gamma\theta_c})/\lambda t \qquad \ldots\ldots(14),$$

where $n=2$ or 3 according to the material. Equation (13) becomes, when θ_c is made very large

$$\lambda t/ds = \sum_{1}^{\infty} (n-1)\,!\,(\beta/n)^n \qquad \ldots\ldots(13.1).$$

This is identical with an equation given by Wagner[4] for a rather different type of breakdown.

Experimental methods. The square law, or constancy of conductivity with respect to field-strength applies to few dielectrics at high field-strengths, and if moisture is present the departures are greater, while the effect of temperature is different at different field-strengths. Accordingly, for simplicity in interpretation of the results, it was necessary to ensure that the thermal resistance outside the specimen should be sufficiently great in comparison with that of the specimen. In addition, precautions had to be taken to ensure that concentrations of stresses and discharges in the ambient medium should not occur, while at the same time satisfactory contact had to be provided so that power factor and permittivity results should be valid, or at least repeatable. The first two of these conditions were fulfilled in one apparatus, figure 2 (*a*), by using rounded edges with oil immersion and a thermal backing of oak discs with cotton-wool packing; and in another apparatus, figure 2 (*b*), by using a very narrow gap between inner plate and guard ring and protecting the outer edge with mycalex, which is arc-resisting. A thermal backing of varnish-paper discs was used. The first apparatus was used in an oil* bath kept at uniform temperature, and the latter in an air thermostat. Contact with cellulose materials, such as ebonite, was made by painting graphite (aquadag) films on the surface; these give satisfactory contact under oil as well as in air. Plain brass electrodes were used for papers which, if flat, give repeatable values on the application of a suitable pressure, though these values cannot be taken as characteristic of the material. The external thermal resistance per specimen was 1·9° C./W. for figure 2 (*a*) and 3·2° C./W. for figure 2 (*b*)†. It was determined by replacing the specimens by flat eureka heaters made so that the energy-inputs under the inner plate and guard ring were proportional to their respective areas.

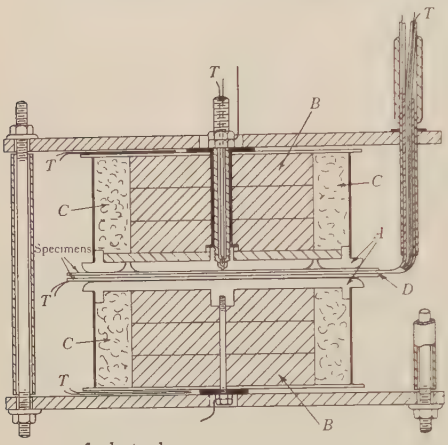
Such a test condenser formed one arm of a Schering bridge. A voltage was applied and maintained until stability of temperature, power factor and permittivity was approached; the voltage was then raised in successive steps until instability

* The oil, Wakefield's superforma, was colourless and the power factor was generally less than 0·001.

† Preliminary tests on varnished cloth were made with a simpler apparatus having a thermal resistance of 5° C./W.

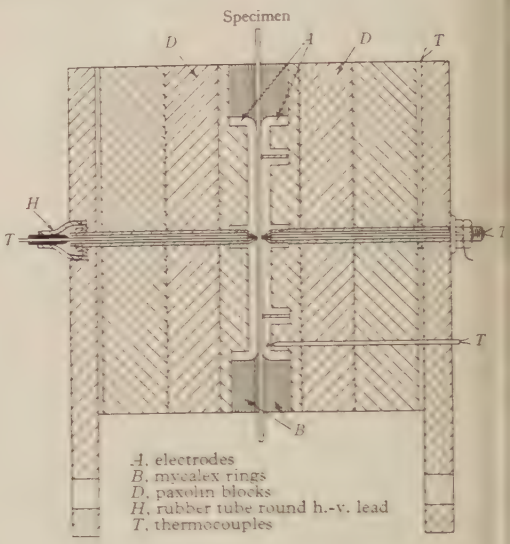
occurred with eventual breakdown, unless breakdown occurred first. A frequency of 50 c./sec. was used throughout.

Separate power-factor and permittivity measurements were made with similar electrodes, but with the thermal packing absent. D.-c. conductivity measurements were made under oil with electrodes similar to those shown in figure 2(*a*), but



A, electrodes
B, thermal insulation of wood blocks
C, cotton wool
D, h.-v. plate
T, thermocouples

(*a*) Oil-immersed apparatus.



(*b*) Apparatus for use in air thermostat.

Figure 2. Main forms of apparatus.

insulated with mica and glass and carefully screened to avoid leakage and stray charges. Short-time tests of electric strength were made with E.R.A. standard brass electrodes in view of the number of samples needed. The upper electrode is $1\frac{1}{2}$ in. in diameter with its edge rounded to a diameter of $\frac{1}{4}$ in., the lower electrode being a disc 3 in. in diameter. In such tests the time for breakdown is observed for a number of different field-strengths and the value for the given time interval is interpolated.

§ 2. PHENOMENA PRECEDING BREAKDOWN: ENERGY-LOSS FORMULAE

A.-c. conductivity phenomena. Cellulose acetate* and cellulose ethyl ether appear to have a species of ordered lattice structure, but behave, when they contain slight traces of moisture, as if two phases are present. The first is associated with a conductivity appreciable at lower temperatures but having a small temperature coefficient; the second has a negligible conductivity at lower temperatures, but a large positive temperature coefficient. This combination forms a well-known case of the Wagner-Maxwell theory of heterogeneous dielectrics and gives a total

* The cellulose acetate contained a plasticizer such as tricresyl-phosphate or triphenol-phosphate, but from tests made with pure films (prepared by A. M. Thomas) the plasticizer appears to have only a minor effect on the electrical properties considered here.

conductivity or specific energy-loss curve which has an inflection with respect to temperature as illustrated in figure 3*. At higher temperatures the conductivity is mainly determined by the second phase and increases linearly with electric stress and exponentially with temperature as shown in figure 4*, and in agreement with equations (8.1) and (8.2). The second phase, which gives the exponential term, is associated with the moisture present, as is shown by the decrease of this term as the moisture is progressively removed. This is illustrated in figures 5 and 6, the laws for which are similar to the equations quoted, namely

$$\sigma = p_0 + aE \exp \gamma (\theta - \theta_0) \text{ for cellulose acetate} \quad \ldots\ldots(15.1),$$

and

$$\sigma = p_0 + aE \exp (\gamma_0 - \gamma E)(\theta - \theta_0) \text{ for cellulose ethyl ether} \quad \ldots\ldots(15.2).$$

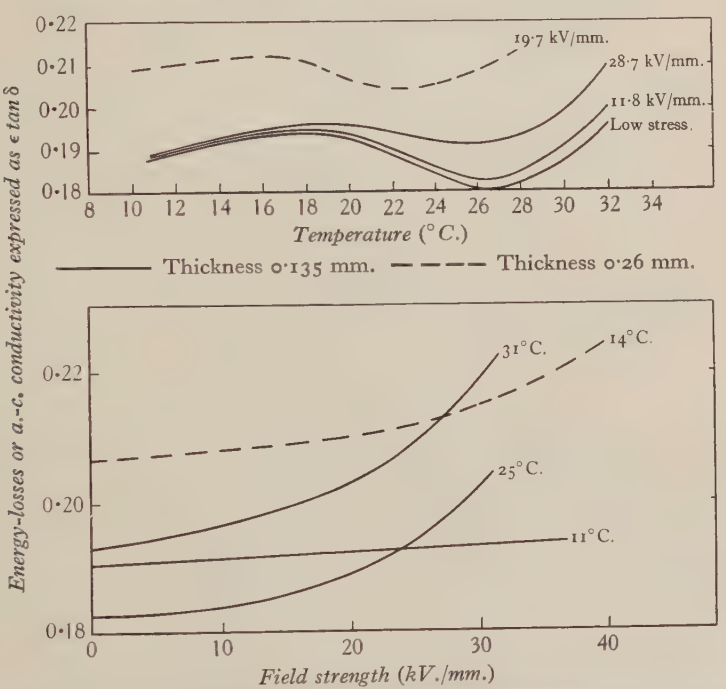


Figure 3. Effect of field-strength and temperature on cellulose-acetate films containing moisture at lower temperatures. (Conditioned 1 hour at 30° C.)

Approximately it may be said that with all these materials equation (10) was obeyed, the departures from this simpler form increasing as the moisture was removed.

A similar conclusion applies to thicker papers (grease-proof paper) and to the thin ebonite illustrated in figure 7, except that in the latter case the variation with electric stress was small so that for ebonite E should be deleted from equation (15.1).

* In the diagrams the product $\epsilon \tan \delta$ is plotted, where ϵ is the permittivity and δ the loss angle. This product is proportional to the conductivity σ since

$$\sigma = f \epsilon \tan \delta \times 10^9 / 2C^2 \text{ mho/cm.},$$

where f is the frequency and C the velocity of light.

Thin dry papers such as condenser tissue gave no appreciable effect of electric stress up to the breakdown range and only a slight variation with temperature. It will be mentioned later that with paper there is the possibility that the electric field may induce ionization in the pores of the material. This possibility does not probably extend to condenser tissue owing to its thinness*. A condenser tissue experimentally impregnated with a paraffin wax showed a variation similar to that denoted by equation (15.2).

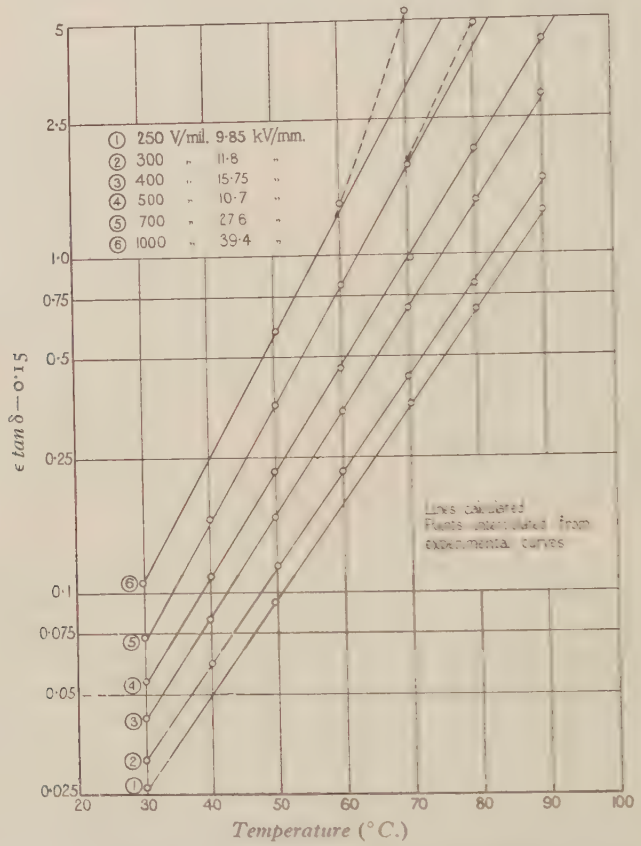


Figure 4. Effect of stress and temperature on specific loss of cellulose acetate, 0·193 mm. thick. (Conditioned 1 hour at 30° C.)

D.-c. conductivity. D.-c. conductivity decreases with time, and two hypotheses are commonly proposed in this connection: the polarization theory which considers the initial conductivity, before polarization has developed, to be the true value, and the absorption theory in which the final conductivity gives the normal conduction. Since from other investigations the polarization potential appears not to exceed a few volts and since the initial values of conductivity were found in the

* Hartshorn and Ward[5] found a sudden increase in the conductivity of condenser tissue at a field of 8·5 kV./mm. The difference may be due to the differences in the form of the electrodes or in the materials employed.

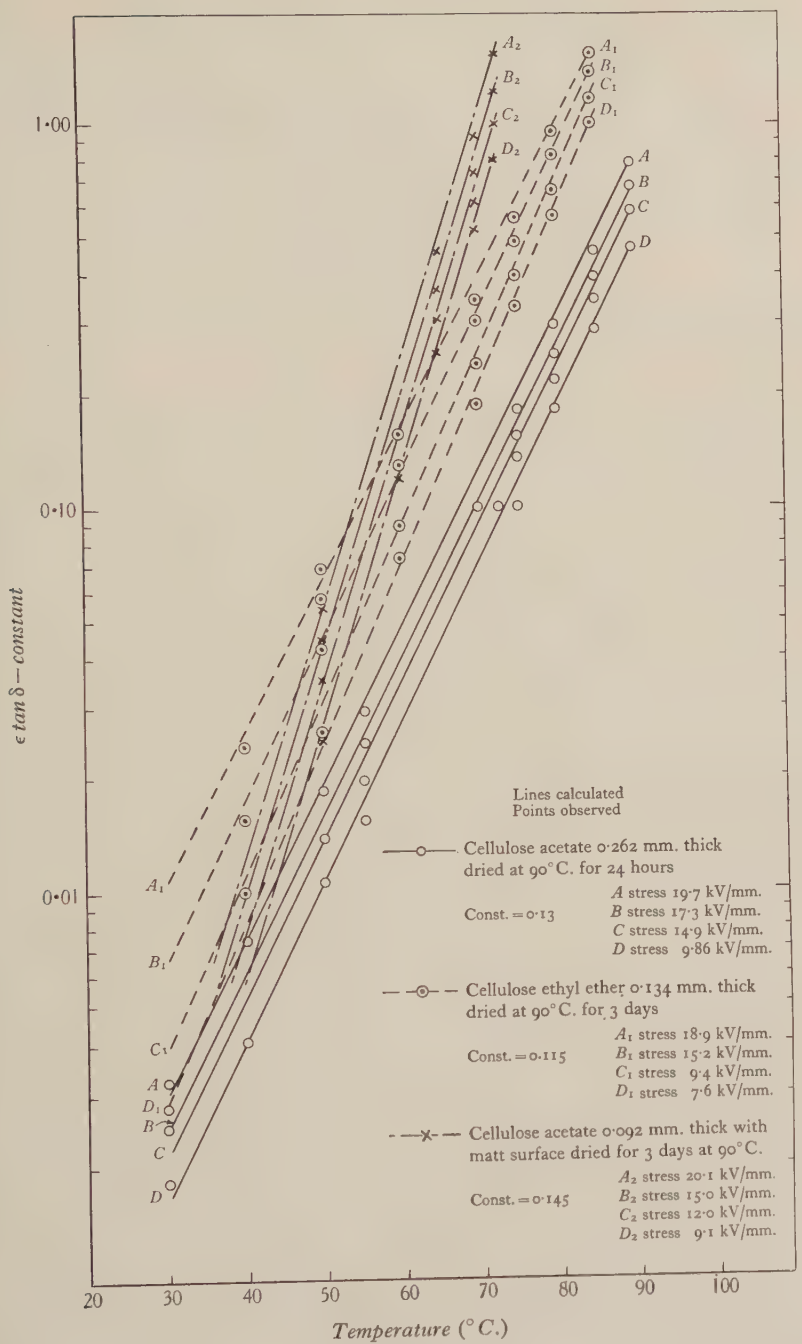


Figure 5. Energy-losses of cellulose compounds with slight traces of moisture.

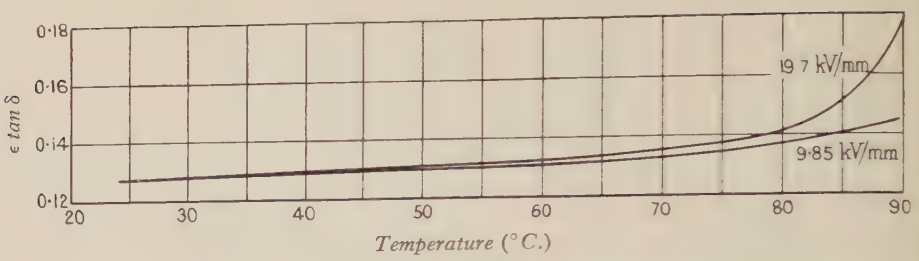


Figure 6. Effect of temperature and stress on the energy-loss of cellulose acetate, 0·265 mm. thick. (Dried for 14 days at 90° C.)

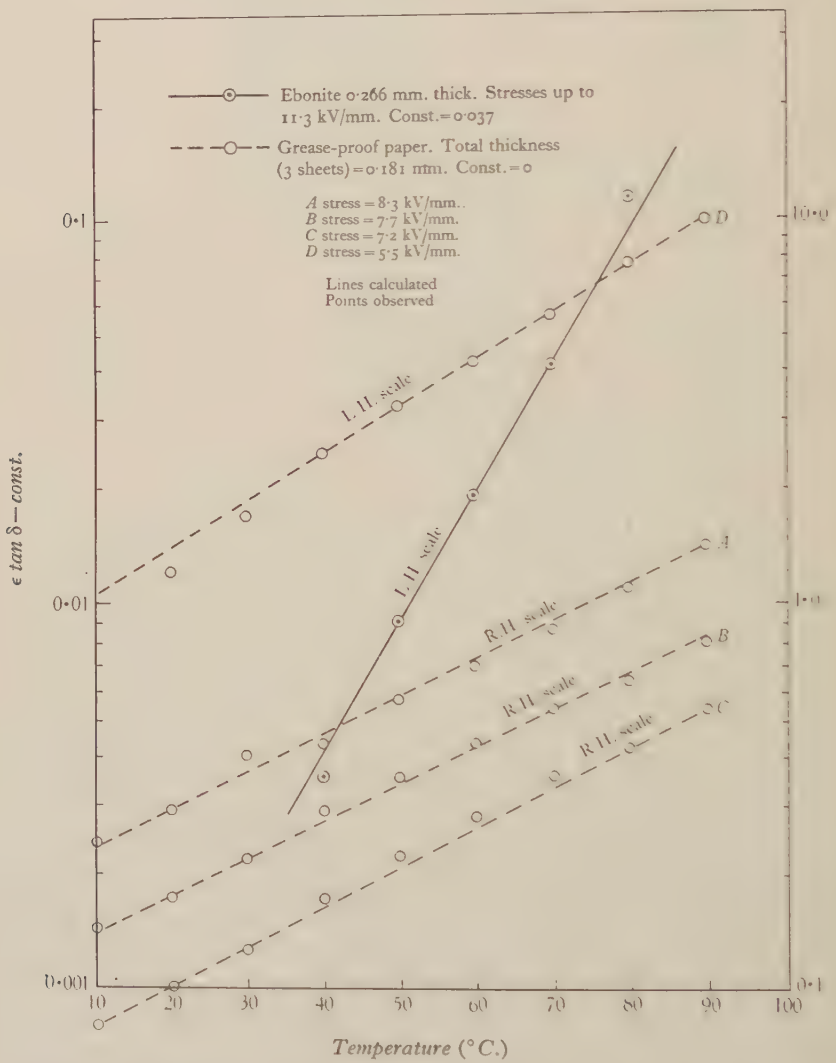


Figure 7. Energy-losses of special thin ebonite and grease-proof paper with slight traces of moisture. (Conditioned at 90° C.)

present tests to be variable and uncertain, while the final values could be repeated over a cycle of electric stress, the absorption theory was adopted and the value for 100 minutes' application of stress was taken as the true conduction. Previous electrification was discharged from a specimen by the application of an alternating field which was gradually reduced to zero.

These conditions being employed it was found that the conductivity of cellulose acetate increased linearly with stress and logarithmically with temperature. The effect of temperature was somewhat complicated. The best expression was found in one fairly extensive series of tests to be as follows:

$$\sigma = Ae^{\gamma_0 \theta} + (\theta - \theta_0)^{\gamma}\, aE \qquad \text{.....(16).}$$

This law is illustrated in figure 8. Thus the d.-c. conductivity also obeys the law

$$\sigma = \sigma_0 + aE \qquad \text{.....(16}a\text{),}$$

but σ_0, in addition to a, is a function of temperature.

As with a.-c., slight traces of moisture have a very great effect; variations of the order of 10^4 to 1 may be produced by varying the time and temperature of conditioning. For example, one specimen after being dried for 1 hour at 30° C. gave, when tested at 90° C., values for σ_0 and a of $5{\cdot}07 \times 10^{-11}$ mho/cm. and $0{\cdot}445 \times 10^{-11}$ mho/cm. per kV./mm. respectively. When the same dielectric was dried for 97 hours at 75° C. the corresponding values for σ_0 and a were $1{\cdot}84 \times 10^{-14}$ mho/cm. and $0{\cdot}14 \times 10^{-14}$ mho/cm. per kV./mm. respectively.

It is, therefore, reasonable to conclude that the large effect of stress and temperature is mainly due to the presence of moisture, although such effects still appear to a limited extent in dried specimens.

Unconditioned specimens show the hysteresis effect observed with ascending and descending electric stress for moisture-absorbing dielectrics. Metastable equilibria with the ambient medium seem to occur which are liable to move suddenly to new equilibrium values as the test is continued. Conditioned specimens are in internal equilibrium and do not show a hysteresis effect. The decay with time of the absorption current seems to be exponential.

Application of the theory of ionization coefficients. Many of the phenomena mentioned can be explained qualitatively on the theory that ions may form others by their motion in an intense field in a manner analogous to ionization by collision in gases, although the precise mechanism cannot be suggested at the moment. It is of interest to consider the quantitative application of this theory and some of the possibilities and difficulties which arise in connexion with it.

For the sake of brevity only one case is taken, namely that in which the ionization per unit path of an ion is μE, where μ is a constant and E the electric field. Let u be the mobility, ϵ the ionic charge, n_0 the volume density of spontaneous ionization, assumed to be constant, n_0' the ionic density produced at the electrodes, and n_0'' the ionic density from other sources. Suppose also that only ions of one sign have an ionization coefficient. Then solving Townsend's continuity equations for ionization by collision in gases, we have, V being the applied e.m.f.,

μ, E

u, ϵ, n_0

n_0'

n_0''

V

$$\sigma = (n_0\, u\epsilon/\mu E)\,(e^{\mu V} - 1) + n_0'\, u\epsilon^{\mu V} + n_0''\, u\epsilon \qquad \text{.....(17).}$$

If only the first term containing E is retained, this reduces to

$$\sigma = \sigma_0 + aE \qquad \text{.....(18),}$$

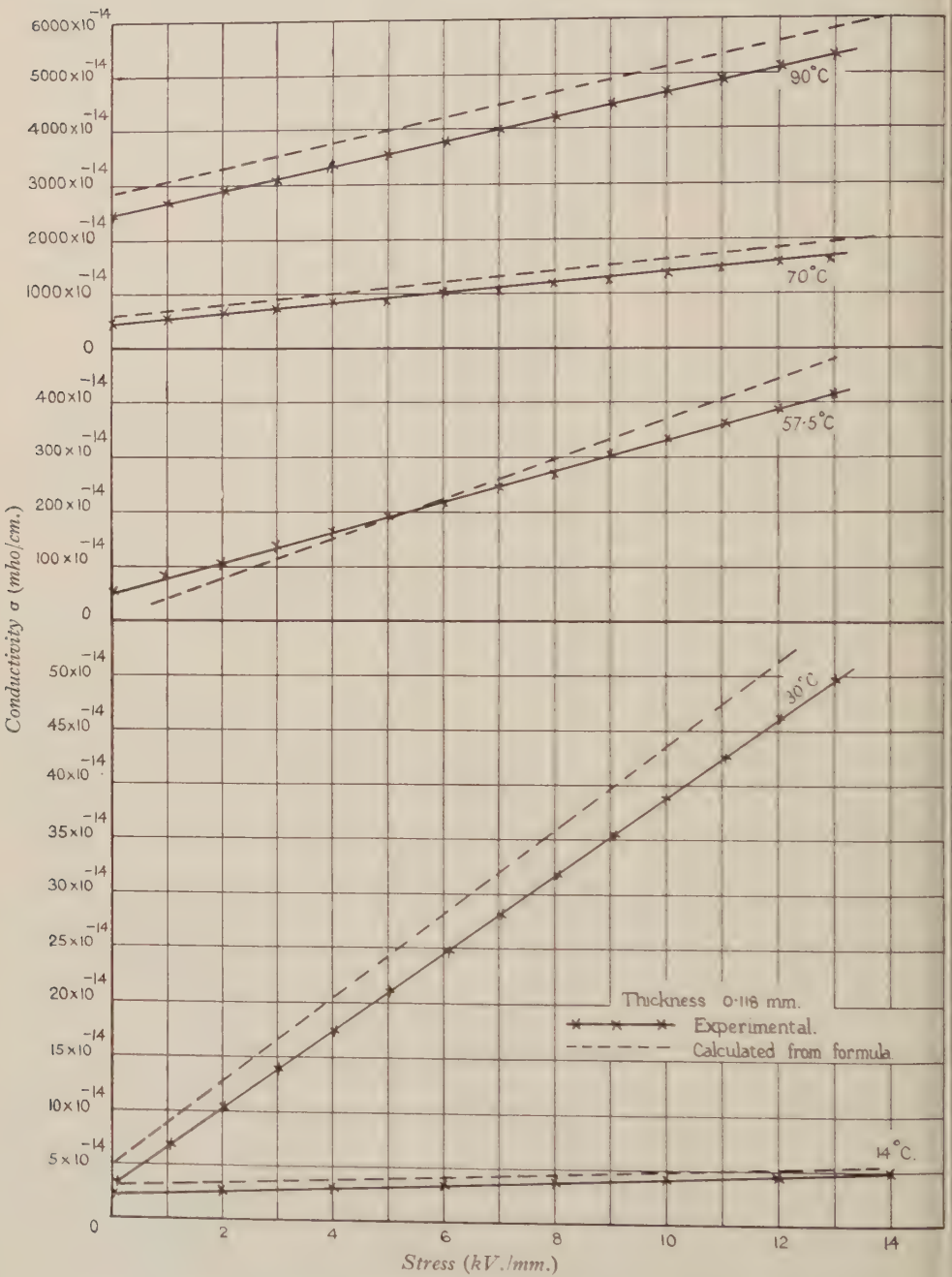


Figure 8. Relation between conductivity, stress and temperature for cellulose acetate containing moisture. (Conditioned 1 hour at 30° C.)

which has been encountered experimentally above. We observe that σ depends on the thickness, but with the materials under consideration other variations with thickness mask such phenomena. It is probable that the third term may be neglected in most dielectrics, particularly if they are moist.

With a.-c. the corresponding conductivity σ' is given by σ'

$$\sigma' = \frac{n_0 u\epsilon}{\mu E}\sqrt{2}.I_1(\mu V\sqrt{2}) + \frac{8n_0 u\epsilon\,\mu V}{\pi E}\sum_0^\infty \frac{(n+1!)^2(2\mu V\sqrt{2})^{2n+1}}{(2n+2)!\,(2n+3)!}$$

$$+ n_0' u\epsilon\left[I_0(\mu V\sqrt{2}) - \frac{1}{\mu V\sqrt{2}} I_1(\mu V\sqrt{2}) + \frac{8}{\pi}\sum_0^\infty \frac{(n+1!)^2(2\mu V\sqrt{2})^{2n+1}}{(2n+1)!\,(2n+3)!}\right] + n_0'' u\epsilon \quad \ldots\ldots(19).$$

E and V are r.-m.-s. values, I_0 and I_1 are Bessel functions of the first kind, with imaginary argument and of zero and first order respectively. The summation terms arise if the d.-c. relations are true for instantaneous currents and e.m.f.s independently of field-direction, i.e. for the symmetrical case. If, however, when the field is reversed, the d.-c. equations hold with E negative, then the summation terms are absent: this is the asymmetrical case. In either case, however, if only the first term is retained, we have I_0, I_1

$$\sigma' = \sigma_0' + a'E \quad \ldots\ldots(20),$$

which has been encountered experimentally above.

We note that experimentally the first term shows extremely wide variations depending on the moisture present. This is to be anticipated theoretically since this term depends on the initial ionization provided by electrolysis of the moisture. With a.-c. there is an additional term, due to absorption, which is fairly constant so that the variations are not very great and cannot be directly correlated with d.-c. values. It may be noted, however, that the difference is less at higher temperatures and higher moisture-contents where the absorption term is of less importance.

The calculation of μ is determined by the ratio a/σ_0 or a'/σ_0' and depends on the relative importance of n_0, n_0' and n_0''. The latter may usually be neglected, and in the cases considered the order of μ is affected in about the ratio of 2:1, according to whether n_0 and n_0' is the determining factor. Since it is only desired to give the order of μ, n_0' is neglected in order to give a basis of calculation, as also are the summation terms with a.-c.* If μ is constant then an ion gives on an average one

* It may be remarked that this type of conduction causes harmonics in the current wave. These have been observed by Gemant[6] when Ohm's law is departed from, but quantitative application cannot be made. If the field is taken as $E\sqrt{2}.\sin\omega t$, then the current i for the case considered is i given by

$$i = b_0 + \sum_1^\infty a_n \sin n\omega t + b_n \cos n\omega t.$$

For the symmetrical case

$$b_n = 0, \quad a_n = (-1)^{\frac{n+1}{2}}\left[I_n(x) + \sum_{k=1}^{k=\infty} (\tfrac{1}{2}x)^{2k}/\Gamma(2k+n+2)\,\Gamma(2k-n+2)\right].$$

For the asymmetrical case

$$4b_0 = x\frac{d}{dx}x^2 I_0(x), \quad b_{2n} = (-1)^n\, 2xI'_{2n}(x), \quad a_n = (-1)^{n+1}\, 2x\, I'_{2n-1}(x),$$

where $x = \mu V\sqrt{2}$, i.e. x is the peak applied potential divided by the ionization potential. Philipoff[7] has given oscillograms showing the development of harmonics during thermal breakdown of the type ending in an explosive discharge.

additional pair of ions in falling through a constant potential-difference of 1 μ. This may be considered as a species of ionization potential and is so defined in table 1.

Table 1

Experimenter	Material	Type of current	Temperature (° C.)	Ionization potential (V.)
	Cellulose acetate:			
Present authors	0·0086 mm., dried at 30° C. 1 hour	d.-c.	51·5	260
		a.-c.	51·5	315
,,	0·0086 mm., dried at 30° C. 1 hour	d.-c.	90	480
		a.-c.	90	190
,,	0·0086 mm., dried at 75° C. 8 hours	d.-c.	90	310
,,	0·0086 mm., dried at 75° C. 73 hours	d.-c.	90	435
,,	0·0086 mm., dried at 75° C. 97 hours	d.-c.	90	550
,,	0·0118 mm., dried at 30° C. 1 hour	d.-c.	14	500
,,	0·0118 mm., dried at 30° C. 1 hour	d.-c.	30	98
,,	0·0118 mm., dried at 30° C. 1 hour	d.-c.	51·5	110
,,	0·0118 mm., dried at 30° C. 1 hour	d.-c.	70	220
,,	0·0118 mm., dried at 30° C. 1 hour	d.-c.	90	370
Hartshorn and Rushton	0·0118 mm., unconditioned	d.-c.	21	180
,,	0·0118 mm., dried at 90° C.	d.-c.	25	130
Present authors	0·0191 mm., dried at 40° C. 1 hour	d.-c.	90	95
,,	0·0191 mm., dried at 30° C. 1 hour	d.-c.	90	120
,,	0·0191 mm., dried at 20° C. 1 hour	d.-c.	90	110
,,	0·0191 mm., dried at 11° C. 1 hour	d.-c.	90	65
,,	0·0191 mm., dried at 30° C. 1 hour	a.-c.	90	210
	Varnished cloth:			
Hartshorn	L 105, dried at 90° C.	50 c./sec.	20	510
,,	L 105, dried at 90° C.	100 c. sec.	20	200
,,	L 105, dried at 90° C.	d.-c.	20	650
,,	L 101, dried at 90° C.	100 c. sec.	20	050
,,	L 101, dried at 90° C.	800 c. sec.	20	200
,,	L 101, dried at 90° C.	d.-c.	20	830

The results indicate that the ionization potential, though variable, does not change characteristically with moisture-content and with type of current, though there may be an increase with temperature. Furthermore the variation, though large on an absolute scale, is small compared with the very wide range of variation of the conductivity in the same conditions. It may be noted that the results obtained with varnished cloth by Hartshorn were fitted to a curve of the form

$$\sigma' = \sigma_0' + a'E^2 + b'E^4 \qquad \ldots\ldots(21),$$

although the last term was not appreciable with cellulose acetate. This result is in favour of what we have called the asymmetrical case.

The linear variation with stress or a variation of the kind indicated by equation (21) has been found by other experimenters such as Poole and Schiller and also in radio-frequency experiments carried out for the E.R.A. In many cases, however, an alternative form

$$\sigma' = \sigma_0' + aE^\gamma \qquad \ldots\ldots(22),$$

proposed by Hartshorn, also agrees with experimental results.

The variation with temperature might be expected to follow the law $e^{-\gamma/T}$ rather than $e^{\gamma\theta}$, where T is the absolute temperature. T

Over small ranges there is not very much difference, but the latter form was found on the average to agree better with experiment.

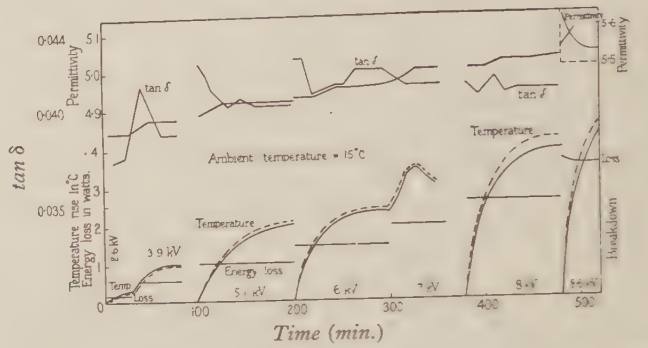
One may conclude that there is a case for the application of the theory of ionization coefficients, since such correlations as are effected by it between the results for a.-c. and d.-c. and between those for other variable conditions represent an advance. Such correlations have hitherto only been possible to some extent for absorption energy-losses and for poor conductors obeying Ohm's law. It is clear, nevertheless, that the phenomena require further analysis in order that various factors may be separated. For instance, with a.-c. the absorption loss must be subtracted while with d.-c. other forms of conduction are present besides those envisaged. The conception of an ionization potential will be considered later in connection with electric strength.

§ 3. PHENOMENA OF INSTABILITY

Breakdown processes. Many materials, particularly if unconditioned, become chemically unstable under prolonged high electric stress, so that although thermal instability may occur the phenomena cannot be predicted from a theory based on stable properties, since the electrical properties do not persist. The materials for which formulae were developed for the stable electrical properties in § 2 were, however, stable in so far that, when they were exposed to prolonged electric stress in the apparatus of figure 2, their electrical properties agreed with the formulae already given, provided the temperature of the specimen, as measured by the thermocouples, had approached fairly close to the final value corresponding to the applied stress. This agreement was also exhibited, to a somewhat lower degree of accuracy, for all electric stresses such that the rate of temperature-rise decreased to a point at which it did not vary greatly during a given measurement, even though subsequently the specimen became unstable under the same electric stress. Further, the energy-losses determined from the temperature-rise when thermally stable, agreed with the electrical measurement of a.-c. conductivity within the accuracy of the thermal determinations which, owing to the low temperature-rises, was about 10 per cent. Thus it may be assumed that substantially all the electrical energy is converted into heat.

Three types of behaviour at breakdown can be distinguished. (i) The temperature-rise, if appreciable, is either stable or increasing slowly at breakdown, which takes the form of an explosive discharge. The conductivity is also stable or else increasing slowly in dependence on the temperature. This type of behaviour was shown by all materials at low temperatures and by dried materials at all the temperatures employed. With ebonite and paper the temperature-rise was very small owing to the very low conductivity. Figure 9 illustrates the behaviour of cellulose acetate containing moisture at an ambient temperature of 15° C. For dry ebonite, paper, and cellulose compounds only a slight temperature-rise and a

slight variation of conductivity occurred right up to breakdown. (ii) The conductivity and temperature become unstable beyond a certain electric stress, but an



Mean temperature rise of l.-v. plates ——. Temperature rise of h.-v. plate -----.

Figure 9. Sustained test. Cellulose acetate, 0·26 mm. thick, containing moisture.

explosive discharge occurs while the conductivity and temperature are still o normal magnitude. This is illustrated in figure 10 for cellulose ethyl ether. (iii) Th

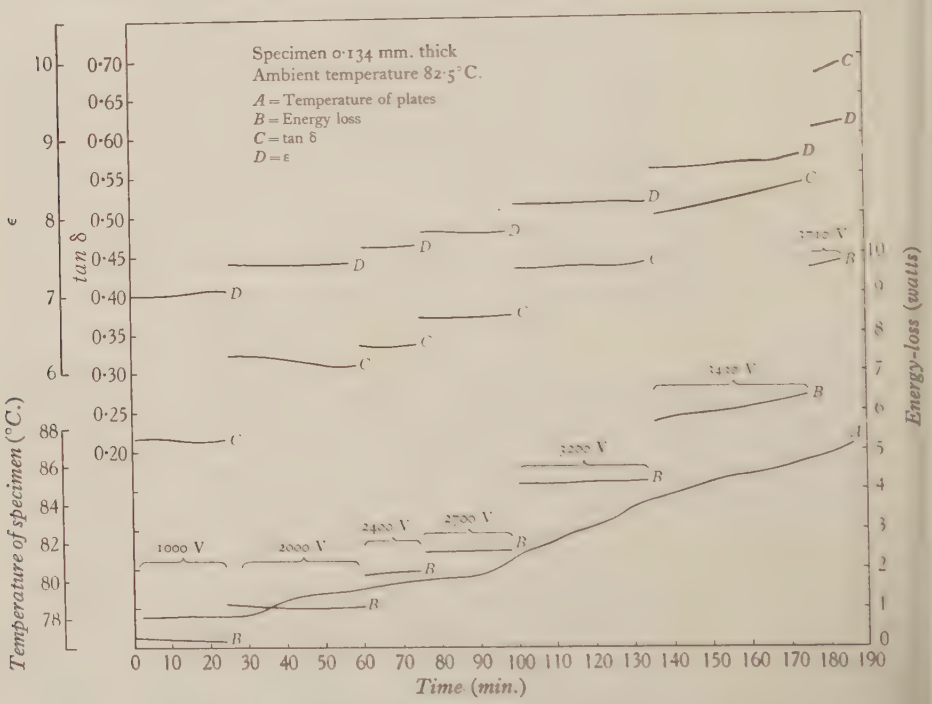


Figure 10. Sustained test on cellulose ethyl ether containing slight traces of moisture.

conductivity and temperature become unstable and the former rises to values suc that insulating properties are lost. The temperature may reach the limit of chemic stability or a discharge may occur. An exaggerated case is shown in figure 11 whi occurred at a high ambient temperature.

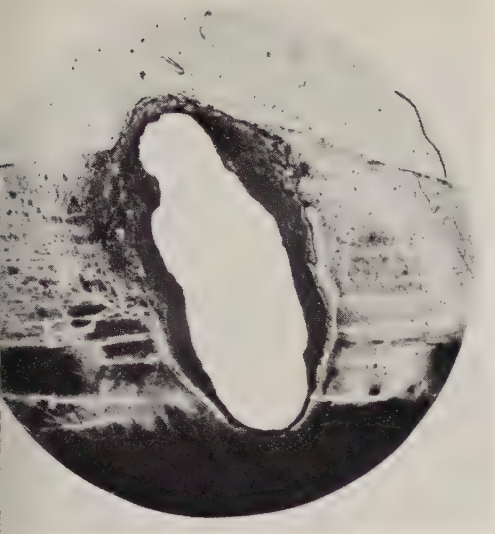


Figure 12 (*b*).

Figure 12 (*c*).

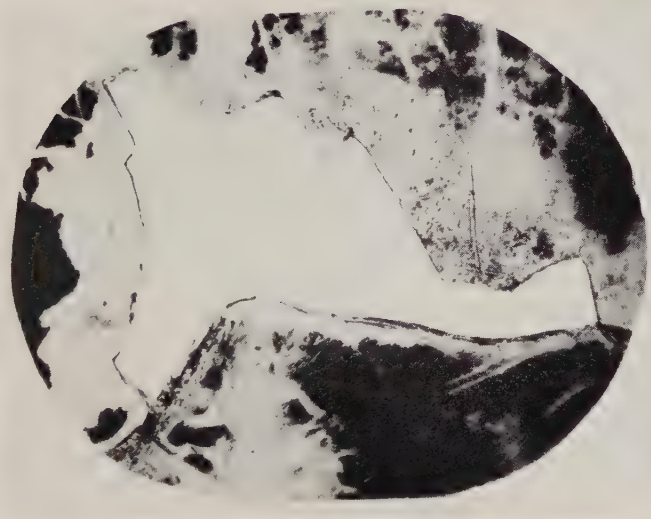


Figure 12 (*e*).

Figure 12 (*d*).

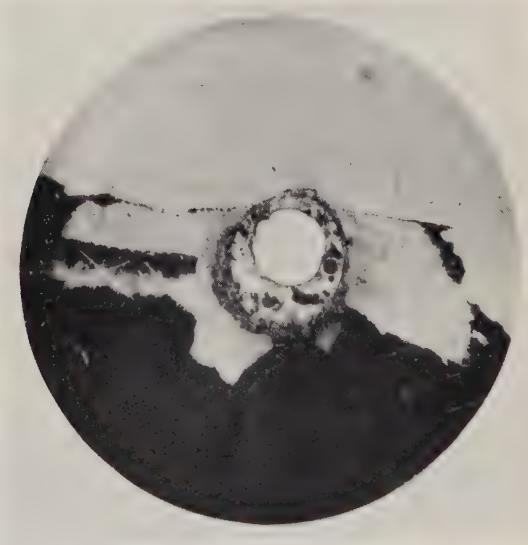


Figure 12 (*a*).

Photomicrographs of punctures.

Figure 12 shows photomicrographs of punctures. (*b*) and (*c*) were taken with air and oil immersion respectively, and indicate that the blackening is due to internal reflection and not charring. (*a*), (*b*) and (*c*) correspond to type (i) above, the discharge in (*a*) being of very short duration. In (*b*) the initial circular puncture spread because the discharge was allowed to continue for a short time. (*d*) and (*e*) correspond to types (ii) and (iii). Similar results were obtained with short-time tests of electric strength; punctures similar to figure 12(*a*) or (*d*) occurred according as thermal instability did not or did take place. The shorter the time of discharge the smaller does the non-thermal puncture become, down to the point where it is

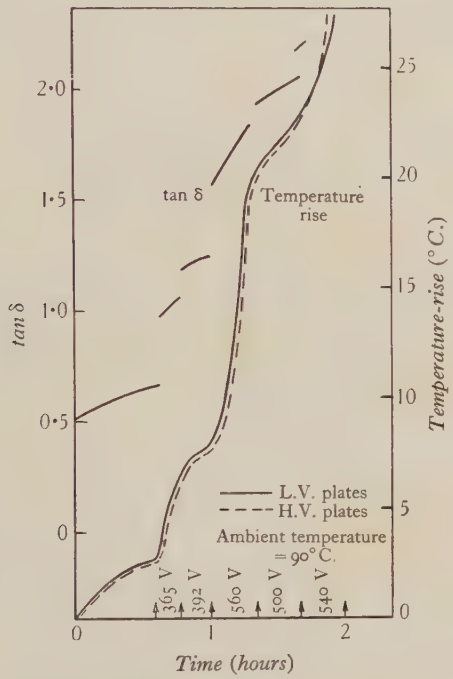


Figure 11. Instability of cellulose acetate 0·05 mm. thick, containing moisture. (Conditioned 1 hour at 30° C.)

scarcely visible. Type (*a*) was often situated at the edge of the electrode, but type (*d*) was nearly always inside. A. M. Thomas[8] has shown that punctures of type (*a*) appear to occur when extraneous discharges are eliminated. In the present instance excessive precautions such as were used by him were not employed. Nevertheless, the appearance is similar.

Variation and limits of thermal electric strength. The theory given in § 1 shows that for sustained tests the thermal electric strength should be calculable and should follow approximately a logarithmic law with respect to ambient temperature, while short-time tests should give curves parallel, on logarithmic paper, to the curve for sustained tests, the height of the curve increasing with decreasing time of stress. Figure 13 illustrates this from experiments on partially dried cellulose acetate. The curves mainly consist of two portions, one of which shows little variation with tem-

perature and refers to lower temperatures, shorter times, and drier and thicker specimens, while the second portion refers to higher temperatures, longer times, and moister and thinner specimens. The second portion obeys the principles of thermal instability set out above and gives quantitative agreement for the sustained tests. The first portion we can relate to a non-thermal type of discharge. Thus the laws of thermal instability are obeyed provided the thermal electric strength is less than the non-thermal electric strength for the conditions of the test. The non-thermal electric strength will be considered later.

The thermal range is associated with instability phenomana of types (ii) and (iii) described above, and with the corresponding micrographical appearance of punctures. The latter applies to short-time as well as sustained tests. The effect of

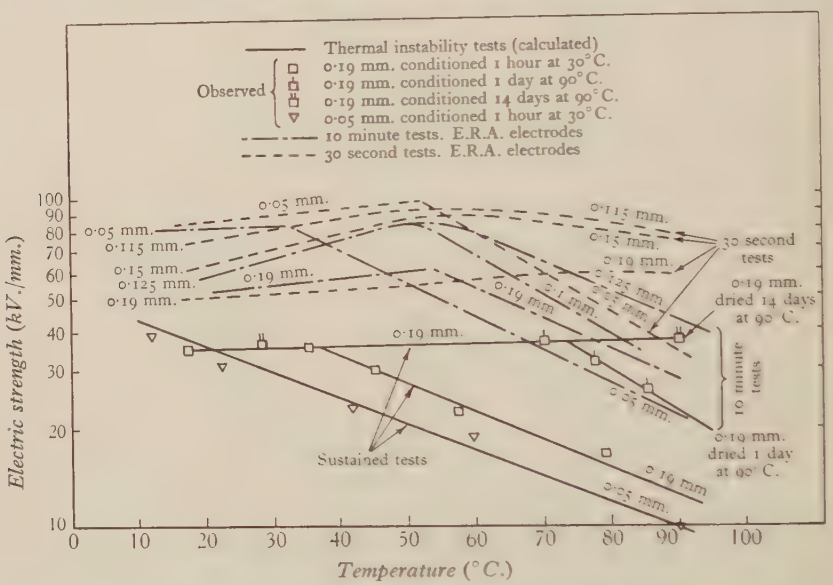

Thermal instability tests (calculated)
0·19 mm. conditioned 1 hour at 30° C.
0·19 mm. conditioned 1 day at 90° C.
0·19 mm. conditioned 14 days at 90° C.
0·05 mm. conditioned 1 hour at 30° C.
Observed
10 minute tests. E.R.A. electrodes
30 second tests. E.R.A. electrodes
Electric strength (kV./mm.)
Temperature (° C.)
30 second tests
0·19 mm. dried 14 days at 90° C.
10 minute tests
0·19 mm. dried 1 day at 90° C.
Sustained tests

Figure 13. Electric strength of cellulose acetate. (Conditioned 1 hour at 30° C. except where the contrary is stated.)

decreasing the time of stress is to move the transition point to higher temperatures. The same result applies to decrease of moisture-content as is shown in figure 13 for tests on films 0·19 mm. thick. Thermal instability did not come within the range of the experiments with the completely dried specimen. Owing to variations in method of manufacture for different thicknesses, the effect of thickness with cellulose acetate films does not follow a simple law. A similar complexity was found with varnished cloth in consequence of increase of energy-loss with the number of layers. With cellulose acetate the thermal electric strength should be inversely proportional to the cube root of the thickness. Actually, owing to the greater energy-loss in thin specimens, the thermal electric strength increases with thickness. With varnished cloth the thermal electric strength should normally be inversely proportional to the square root of the thickness, but actually it decreased more rapidly than this owing to an increase of energy-loss with number of layers.

We have neglected the maximum and minimum of energy-loss which occurs low 30° C.—the so-called *V* curve. It has no influence on the present results and ermal breakdown can only be realized at higher temperatures. A detailed dis- ssion has been given elsewhere and it need only be remarked that a slightly omalous electric-strength curve is produced, the anomalous curve giving appar- tly lower values of electric strength than the normal curve.

Cellulose acetate with a matt surface and cellulose ethyl ether conditioned at ° C. gave results similar to those given by cellulose acetate, the transition tem- eratures being different owing to differences in the energy-losses and to a smaller egree in the non-thermal electric strength. Thermal breakdown occurred above ° C., in agreement with theory for the matted material with sustained tests. For llulose ethyl ether, the theoretical transition temperature was 54° C. on the ssumption that the non-thermal electric strength was 39 kV./mm. The specimens ere not, however, very uniform in this respect and in some sustained tests a non- ermal electric strength of 29 kV./mm. was indicated. In such circumstances the eoretical transition temperature was 84·5° C. The majority of the tests, however, nfirmed the theoretical value of 54° C.

With dried ebonite the theoretical transition temperatures in sustained tests fell utside the range of the experiments, being 94, 97·5 and 101° C. with thicknesses of ·52, 0·385 and 0·275 mm. respectively. Owing to the low non-thermal electric rength of dried grease-proof paper, the transition temperature is very high, being 06° C. for 3 sheets and 143·5° C. for 15 sheets. Condenser tissue paper when dry very slightly affected by stress or temperature up to the non-thermal electric rength, and accordingly thermal breakdown cannot occur except with very great nicknesses where the thermal resistance of the paper itself is much greater. In all ese cases it was verified that breakdown through thermal instability did not occur t the lower experimental temperatures.

Non-thermal electric strength. It is of interest to consider the nature of the lectric strength which limits the application of thermal laws, although this in- estigation did not form part of the original programme and is being pursued by ther means. There is an initial difficulty in determining the influence of edge ffects and extraneous discharges. According to A. M. Thomas it is necessary to tilize a penetrating ambient medium of high electric strength, such as xylene nder pressure, to eliminate these effects, and in such circumstances the electric trength of, for example, cellulose-acetate film is three or more times the values here bserved. On the other hand the type of puncture resembles the type which Thomas considers characteristic of the elimination of edge effects. Figure 14 shows he effect of thickness for non-thermal and thermal breakdown with cellulose cetate. The former follows the parabolic law, the thickness being inversely pro- portional to the square of the average electric strength; this is the law arising from constant limiting maximum field-strength applied to the field between two thin iscs. The effect of introducing a graphite contact film is to increase the electric trength by about 10 per cent. In the sustained tests breakdown occurred more ften within the guard ring than at the edge, following a similar law. With ebonite,

breakdowns fell into two groups according to whether the puncture was at the edi of the electrode or within it. In both groups the electric strength was independe of thickness. Tests on different numbers of layers of paper, both tissue and greas proof, gave a constant electric strength. It is not clear, therefore, how far the

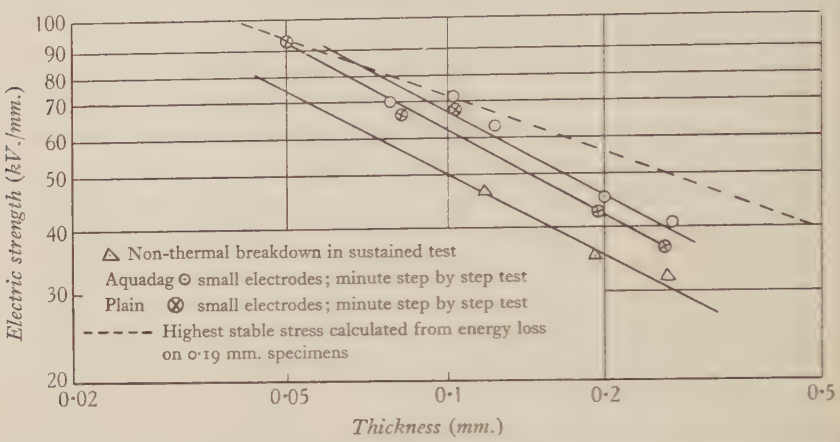


Figure 14. Electric strength of cellulose-acetate films of various thicknesses at 20° C.

values are characteristic of the material or whether the field-distribution of ambi medium exert an important effect, and investigations are proceeding with resp to these questions.

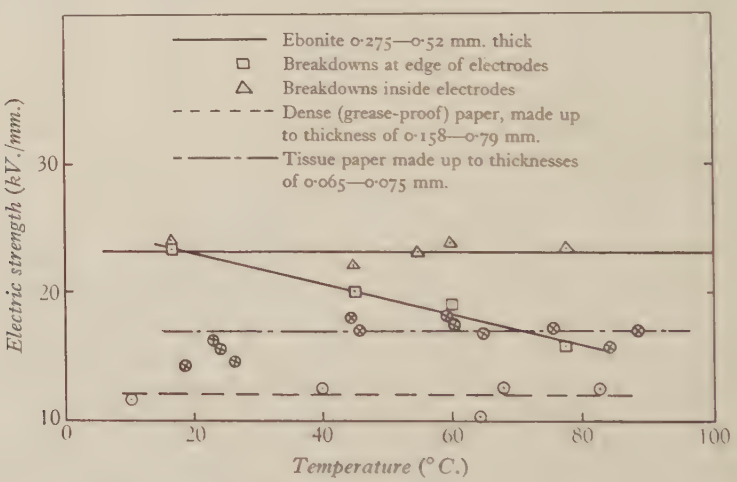


Figure 15. Non-thermal breakdown in sustained tests.

From figure 13 it is seen that temperature has not a very great effect on the n thermal electric strength of cellulose-acetate films. A slight increase with te perature appears possible, but usually it is not great in comparison with inherent variability of this quantity. Figure 15 shows that for ebonite the elec strength for internal breakdowns is constant, while for edge breakdowns it decre

ith temperature. The results of sustained tests on grease-proof paper and condenser ssue paper were also unaffected by temperature, but short-time tests gave some- hat different results. The electric strength of condenser tissue was constant, hile the electric strength of grease-proof paper decreased slightly with temperature hether tested in air, figure 16, or in oil, figure 17, although impregnation with oil ncreases the electric strength considerably to approximately the same value as hen the tissue paper is impregnated with paraffin wax. The electric strength of a ayer of air of the same thickness as a single sheet of tissue paper is of the order of

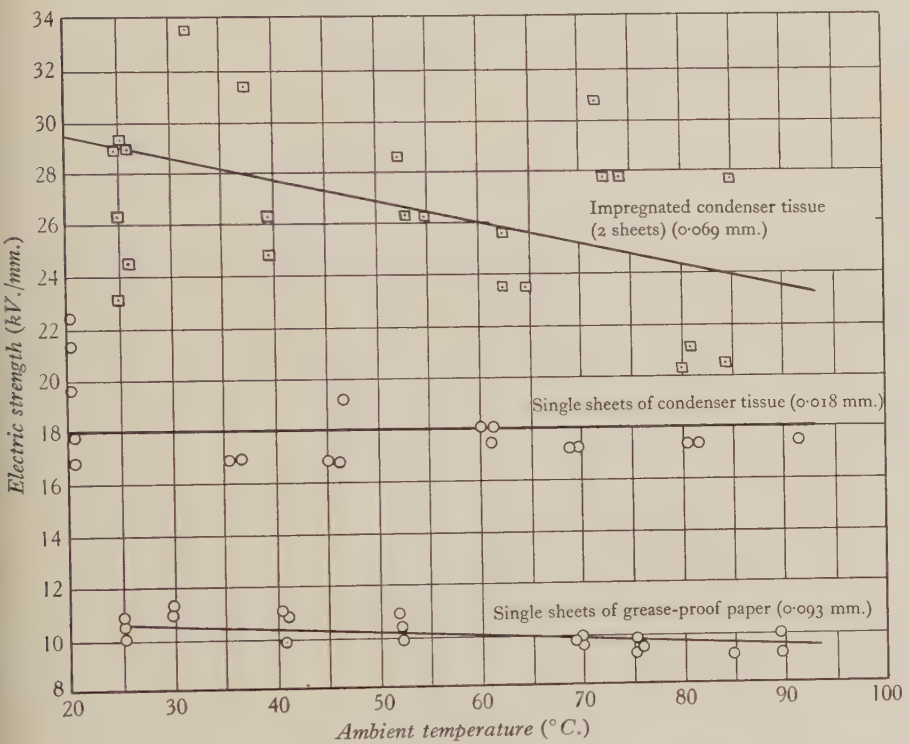


Figure 16. Short-time electric strength of paper in air. (Total time 1 minute.)

30 to 40 kV./mm., while for a single sheet of grease-proof paper, it would be about 10 kV./mm. Thus there seems reason to believe that the linear decrease with temperature is probably associated with the intervention of the ambient medium at edges or at interstices in the structure. The effect is difficult to separate from the inherent variations.

The non-thermal electric strength is not much affected by small quantities of moisture, as for example in the experiments on cellulose acetate films. The effect of duration of test was not studied in detail except that it was noted, as is well-known, that the electric strength decreases with duration of test. This effect is, for durations greater than 10 seconds, much less than the corresponding effect on the thermal electric strength. This finding is in agreement with more recent work on dielectrics

such as that of Inge and Walther, and it may be concluded that the non-therma electric strength of dielectrics does not normally vary greatly with duration o test in the absence of strong external ionization in the ambient medium.

It is also to be noted that corrugation of the surface of cellulose acetate, as i matt films, had no appreciable effect on the electric strength. This is to be expecte since the electric strength of fine channels when filled with air or oil would b greater than that of the solid material.

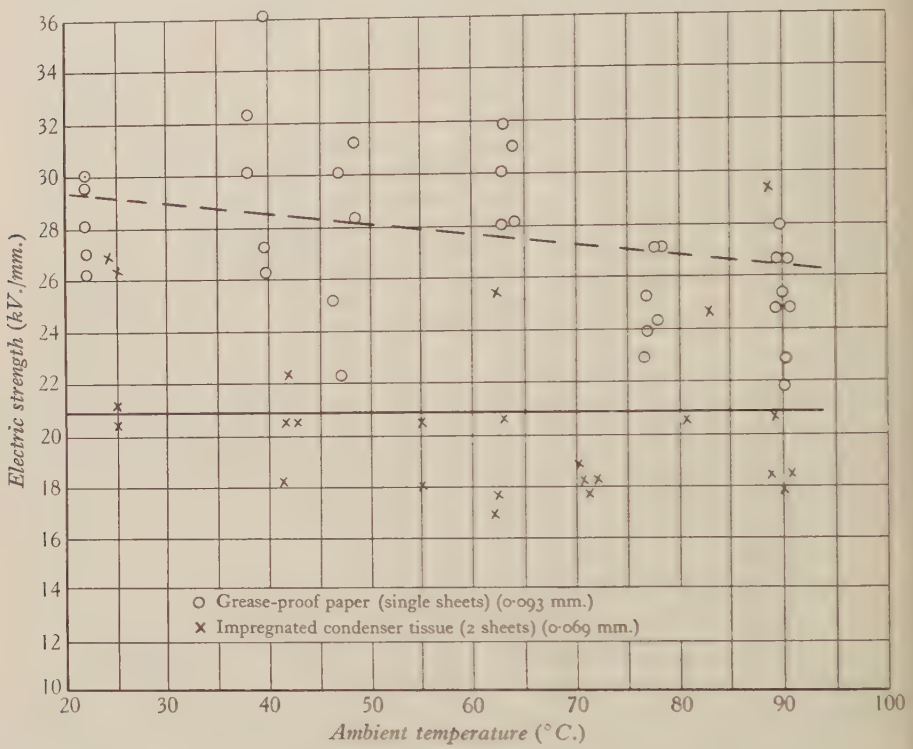


Figure 17. Short-time electric strength of paper immersed in transformer oil. (Total time 1 minute.)

Tests made on paper with continuous voltages gave the same results as thos made under the same conditions with alternating voltages having peak values equ to the values of the continuous voltages.

Modern views of intrinsic electric strength. Although the present paper dea only indirectly with the intrinsic electric strength, which enters as an upper limi it is of interest to consider recent theories dealing with this subject since it appea likely that two or more groups of effects are involved.

Joffe[9] arrives at ionization potentials of the order of 100 or 200 volts, althoug he finds a decrease with field-strength not observed by the present authors. Suc values, and those given in this paper, fit in with his theory of transition from ion breakdown (about 10^6 V./cm.) to structural breakdown (10^8 V./cm.) from thick

esses of 10^{-3} cm. to those of 10^{-2} cm., since the voltage across the specimen has eached the minimum for ionization and cannot fall further until another type of ailure occurs. This, however, is an over-simplification. Ionization can occur at mall thicknesses at lower voltages than are possible with these values of ionization otential: ionization by ions of one sign does not become unstable without the ntervention of some other factor, such as ionization by ions of opposite sign: ionic nergies calculated from structure are very much less than these ionization potentials; nd finally some doubt attached to Joffe's electric-strength values of 10^8 V./cm. Examination of actual discharge currents also supports the view that the high onization potentials are not physically real, but are statistical and average effects ue to energy-dissipation without ionization; and although an ionic form of break-own probably exists, it is related to current-density phenomena and not directly o the structure.

Calculations of ionization potentials of polar molecules usually give low values[9]. Wideroe[10] arrived at ionic energies from the heats of sublimation and dissociation or rocksalt, giving 2·8 electron-volts for the sodium ion and 11·6 for the chlorine on. Hippel[11] indicates a proportionality between the voltage-drop in a lattice lement at breakdown and the energy-difference in electron-volts corresponding o the Reststrahlen of different crystals. His view is that an electron raised to the ighest level of an incomplete energy band by the electric field may, by falling to a ower level in the band, yield energy to an electron in a lower complete band, so hat two free electrons are produced. The ionization potentials so deduced from the voltage-drop per unit element would be only of the order of a fraction of a volt, the igures for electric strength being between 10^5 and 10^6 V./cm.

A calculation has recently been made by Zener[12] of the rate of escape of electrons rom a complete band through an energy-difference into an incomplete band under he action of an electric field. He finds that this rate of escape is so rapidly accele-ated as the field increases beyond, say, the order of 10^6 V./cm., as to constitute reakdown, assuming a lattice constant of the order of 3×10^{-8} cm. and an energy-difference of, say, 2 V. If we take the equations for semi-conduction in crystals worked out by Wilson[13] on the same basis of closed and open groups of electrons in solid lattices and then consider that a species of Stark effect occurs giving new intermediate energy-levels and thus inducing metallic conduction, we find that if the same order of constant occurs as in gases, conduction in a field of the order of 10^6 V./cm. should occur if the energy-level displacement is to be of the order of 0·01 V., which is about the order of energy-differences exhibited with some semi-conductors which are insulators at low temperatures. This view however leads to an interaction of temperature and field-strength giving a term of the form $\exp(\lambda E/kT)$ in the conductivity. Thus the lower the temperature the greater the relative effect of stress, so that although an exponential variation of conductivity with stress is very frequent, it can hardly be ascribed to such a mechanism since the corresponding temperature effect has not been observed.

At present the general conclusion appears to be that the final limiting electric strength is probably nearer 10^6 or 10^7 V./cm. than the 10^8 V./cm. given by Joffe,

and is associated with energy-differences possibly as low as hundredths of a volt but probably not greater than a few volts. This structural breakdown should be sudden in character giving a phase of electronic conduction, at any rate initially which is unlikely to be related to the conduction preceding breakdown or to be greatly affected by time or temperature over a wide range. An ionic form of breakdown, comprising a kind of electrolytic decomposition with induced ionization may occur at lower field-strengths and will be related to the conduction immediately preceding breakdown. The ionic type of electric strength should show time effects at longer times than the structural type but should not necessarily be greatly affected by temperature.

In the present tests the times were too long to enable such distinctions to be made experimentally; the breakdown might have been of either type. In addition it is not yet completely known how to avoid external effects, imperfections in the material, or the effect of a mosaic structure if present. On the other hand the fact that electric strengths of 10^6 V. cm. and above can sometimes be realized with suitable precautions with many materials indicates that these difficulties may not be so great in the electrical problem as in the corresponding mechanical problem. Work is proceeding on the non-thermal type of breakdown at shorter times to attempt a further separation of breakdown types.

§ 4. CONCLUSIONS

(i) Provided that a sufficiently wide range of external conditions is available a dielectric will exhibit at least two types of breakdown—a thermal and a non-thermal type. In the case of thermal breakdown the electric strength corresponding to a long time of stress can be calculated with adequate accuracy from the conductivity under stable conditions, when the external thermal conditions and also the thermal properties of the dielectric are known. Provided the electric strength so computed is less than the non-thermal electric strength, the calculation will be valid whether breakdown takes the form of a progressive loss of insulating properties or the form of a sudden explosive discharge. The thermal electric strength depends on the electrical processes preceding breakdown only in so far as these processes contribute to the generation of heat in the dielectric. Joint electrical and thermal effects of the type suggested by Rogowski and others were definitely absent in the present work.

(ii) If a stronger field is applied so that breakdown takes place in a short time the theory of thermal instability may be extended to give the electric stress corresponding to this time of application. The quantitative agreement between theory and experiment is not then so accurate, but if the variation of electric strength with time is known for a given temperature the variation at other temperatures can be theoretically deduced therefrom with sufficient accuracy.

(iii) A transition temperature exists for any given time of application of electric stress and given external conditions. Below this temperature non-thermal breakdown occurs, above it the electric strength obeys the theory of thermal instability. The range of application of the latter theory is accordingly confined

rgely to higher temperatures, thicker specimens, poorer dielectrics, longer times f application, or electrode arrangements of higher thermal resistance.

(iv) Although there is evidence from the phenomenon of a.-c. and d.-c. conuctivity to the effect that ionic effects occur, and although ionization potentials hich give a fairly simple account of the phenomena can be deduced, yet it is oubtful whether these stable processes bear a direct relation to any of the types f non-thermal breakdown. The experimental methods employed showed no vidence of a factor which would be capable of producing ionic instability and also etermine the conductivity at lower field strengths.

(v) The ionization potentials deduced are much greater than those which can e obtained from considerations of solid structure, and they have not the same hysical significance as the corresponding quantities in gases. On the other hand he electronic theory of intrinsic breakdown does not appear to be capable of xplaining the stable phenomena.

(vi) The electronic form of non-thermal breakdown should be independent of emperature, thickness of dielectric, time of application of stress, and humidity. he same appears to be true of the ionic form, except that this form is affected to ome extent by the time of application of stress.

§ 5. ACKNOWLEDGMENT

The authors wish to thank Prof. MacGregor-Morris, in whose laboratory the xperimental work was carried out, for the facilities provided and for his help uring the investigation.

REFERENCES

(1) *Report of the British Electrical and Allied Industries Research Association*, Ref. L/T 69.
(2) WHITEHEAD. *Dielectric Phenomena*, **3** (Benn).
(3) FOCK. *Arch. Elektrotech.* **19**, 71 (1927).
(4) WAGNER. *J. Amer. Inst. elect. Engrs*, **41**, 283; *S.B. Akad. Wiss. Wien*, **29**, 438 (1922).
(5) HARTSHORN and WARD. *Report of the British Electrical and Allied Industries Research Association*, Ref. L/T 48.
(6) GEMANT. *Arch. Elektrotech.* **23**, pt. 6, p. 683.
(7) PHILIPOFF. *Z. tech. Phys.* **14**, No. 1, p. 21 (1933).
(8) A. M. THOMAS. *Report of the British Electrical and Allied Industries Research Association*, Ref. L/T 57.
(9) JOFFE, see (2).
(10) WIDEROE. *Arch. Elektrotech.* **26**, pt. 9, p. 626 (1932).
(11) HIPPEL. *Z. Phys.* **67**, 707; **68**, 309; **75**, 145.
(12) ZENER. *Proc. roy. Soc.* A, **145**, 523 (1934).
(13) WILSON. *Proc. roy. Soc.* A, **133**, 458 (1931); **134**, 277 (1931).
(14) FOWLER. *Proc. roy. Soc.* A, **140**, 505 (1933).

DEMONSTRATION

THE USE OF A LIQUID SURFACE CARRYING RIPPLES AS A DIFFRACTION GRATING

Given by R. C. BROWN, Ph.D., B.Sc. Lecturer in Physics, University of London, University College.

Demonstration given on June 7, 1935.

IT is well known that from measurements of the wave-lengths of ripples of known frequency the surface tension of the liquid upon which the waves are travelling can be calculated. The most usual method of arriving at the wave-length is that of direct measurement, either of progressive waves viewed stroboscopically or of stationary waves.

Kalähne[1] showed that it was possible to use a rippling surface as a diffraction grating and by measuring the angular separation of the spectra he was able to calculate the wave-length of his ripples, i.e. the grating-space. Stationary ripples were used but, owing to the enormous disparity between the velocity of the ripples and that of light, progressive waves will provide an equally good grating.

The following is a description of a very simple experiment designed to demonstrate the spectra produced. In figure 1, *C* and *T* represent the collimator and telescope respectively of a small spectrometer. They are adjusted for parallel light

Figure 1.

and arranged so that an image of the slit is seen in the telescope by reflection at almost grazing incidence from the surface of water contained in a half-plate developing dish. *D* represents the dipper consisting of a small glass plate attached to the diaphragm of a loud-speaker unit so that when the loud-speaker is actuated by a small valve oscillator the dipper vibrates vertically while keeping parallel to its own plane.

The position of the dipper should be such that no reflection of light takes place from portions of the water-surface which are near enough to it to be curved as a result of capillarity. In order to effect this adjustment, it is convenient to restrict the light leaving the collimator by means of a small piece of card.

If the bench on which the experiment is being performed is liable to vibration the dish may be held in a sling of rubber tubing in the way in which microphones are suspended. It may be necessary also to place a thickness of sponge rubber between the bench and the stand which carries the loud-speaker unit and dipper, for irregular disturbances originate from the dipper if this is not done.

The production of standing waves by reflection from the sides of the dish gives se to a continuous background of illumination and detracts from the clarity of the ectra. This can be avoided by using small amplitudes.

Good diffracted spectra are seen in the telescope when white light is used and hen the frequency of the ripples is between 50 and 500 c./sec. Above this range it difficult to produce waves of large enough amplitude to give any observable effects.

So far the author has not adopted this somewhat indirect method of measuring pple wave-lengths, since it would not be superior to the direct method of measure-ent except at high frequencies, i.e., with wave-lengths too short to be measured ccurately on a divided scale. For frequencies below 500 c./sec. wave-lengths are rge enough to be measured directly with sufficient accuracy.

REFERENCE

(1) KALÄHNE. *Ann. der Phys.* **7**, 440 (1902).

REVIEWS OF BOOKS

The Principles of Quantum Mechanics, by P. A. M. DIRAC, Fellow of St John's College, Cambridge. Second edition. Pp. xi+300. (Humphrey Milford: The Oxford University Press, 1935.) 17*s*. 6*d*. net.

The first edition of this book, which was published in 1930, marked a new epoch in the development of the quantum theory. In this second edition considerable changes have been made in the method of presentation, and a certain amount of additional matter has been included.

The earlier chapters, which built up the general theory of the representation of states by vectors and of observables by linear operators, have been almost completely rewritten and the method of presentation has been somewhat simplified. The theory nevertheless remains highly abstract in character and makes no concession to the physicist who wishes for an objective and realist interpretation of the fundamental axioms of this subject.

In the chapter on the equations of motion Dirac has incorporated some of his recent work on the form taken by the principle of least action in quantum theory, together with a section, which is all too brief, on the theory of the statistical ensemble.

The chapter on the relativistic theory of the electron includes Schrödinger's exact solution of the equations of motion of a free electron, together with the theory of the positron, the existence of which, it will be remembered, was predicted by Dirac over a year before it was discovered experimentally.

The last chapter in the book is entirely new and is devoted to the fascinating subject of field theory, due to Jordan and Pauli, Heisenberg and Born. The interest of this particular development is that it is concerned with the principal problem which so far has defied solution by the methods of the quantum theory. Indeed it appears that some essentially new principle is required to deal with the difficulties of micro-microphysics on the scale of electronic and nuclear dimensions.

The book is beautifully printed and produced by the Oxford Press. The reviewer has discovered only one misprint, namely on page 74, line 4, where an accent is missing from the δ symbol.

G. T.

The Measurement of Inductance, Capacitance, and Frequency, by ALBERT CAMPBELL and ERNEST C. CHILDS. Pp. xxiv+488. (London: Macmillan and Co.) 30*s*. net.

It is a fortunate thing that Mr Albert Campbell has been able to tear himself away from his laboratory for a sufficient time to enable him, with the assistance of Mr Childs, to write this book. Probably no man living has had greater experience in the matters of which it treats than he, and in this work he succeeds in handing on some of the fruits of that experience to other workers.

The book may perhaps be said to be somewhat unbalanced inasmuch as the most modern work at radio frequencies is not presented in quite the same comprehensive and masterly manner as, say, vibration galvanometers, the construction, calculation and measurement of inductors, and condensers, the alternating-current potentiometer, and frequency-meters. But it must be a difficult matter to decide on the exact scope of such a work and incidentally to give it an adequate title. It gives much valuable information, not easily accessible elsewhere, on general electrical measurements. Almost the whole field of alternating-current measurements may be said to have been covered, including measurements at low, audio and radio frequencies, but there is no doubt that it will be found most

useful by the man who is concerned mainly with audio-frequency measurements. Many interesting details concerning the history of the subject are touched upon. For example, many physicists must have wondered who coined the terms inductometer and variometer, and what is meant by them. Mr Campbell is a safe guide on such matters. The book should be noted by all physicists who are concerned with measurements involving the use of alternating currents, including work on dielectrics, magnetic materials, and electrolytes, as well as the more obvious measurements on condensers, inductances, and laboratory standards generally. The book is very clearly and concisely written, conveniently arranged and indexed, and in every way well adapted to its purpose. L. H.

Intermediate Physics, by C. J. Smith, Ph.D., M.Sc., A.R.C.S. Second edition. Pp. xii+900 with illustrations and diagrams. (London: E. Arnold and Co.) 16*s*. Also published in parts.

The original issue of this book was criticized for a certain lack of balance, but this has been corrected in this new edition. The section on electricity and magnetism has been thoroughly revised and amplified, and the opportunity has been taken of making small improvements and additions throughout.

The text now provides a very suitable background for the usual type of intermediate course, especially for the abler students. It will be particularly useful for those preparing for more advanced work to come, or aiming at scholarships. Ostensibly no previous knowledge of the subject is assumed, but the book does not appear to be suitable for an absolute beginner. It is mainly experimental in outlook, but does not profess to provide the detail which is looked for in a practical text-book. It pays little attention to the historical aspect of the subject—which is a pity, since this is too often under-valued.

Those familiar with the first edition will be interested to learn that the book can now be obtained in separate parts, five in number, the dearest being that on electricity and magnetism at 6*s*. Short mathematical tables are included, but numerical data of a physical nature are still inadequately provided, and the student must glean what information on magnitude he can from chance numbers in the problems. The theory of dimensions occupies a more prominent place than in the previous edition, and there is more scope for the student with good mathematical equipment. The diagrams are even better than before.

The general treatment is very sound, and pleasurably free from fallacies and doubtful "simplifications". This does not mean that there are no errors at all. Indeed, the author's ideas on laws, physical and penal, and on the physical meaning of negative magnification, for example, need revision; and here and there an example or a phrase is used before the way has been prepared for it in the text. But the discriminating teacher will easily remedy these defects, and will probably find the book no worse in this respect than most others of its size and scope. It is certainly very good value for money, and can be strongly recommended. A. J. W.

Through the Weather House, by R. A. Watson-Watt. Pp. xi+192. (London: Peter Davies, Ltd.). 7*s*. 6*d*. net.

Probably many Fellows of the Physical Society listened with no less pleasure than the general public to the popular series of talks on meteorology broadcast by Mr Watson Watt early in 1934. These lectures are now published in revised and expanded form under the title *Through the Weather House*, recalling the picturesque imagery with which Mr Watt adorned his fascinating tale of our atmosphere, its beauties, its wonders, and its problems. The book deserves and will assuredly gain a wide sale. S. C.

Fine Structure in Line Spectra and Nuclear Spin, by S. TOLANSKY. Pp. viii + 112. (Methuen's Monographs on Physical Subjects. London: Methuen & Co., 1935. 3*s*. net.

In recent years optical spectroscopy has added much to our knowledge of the atomic nucleus. The magnitudes of the angular momenta and magnetic moments of a large number of nuclei have been derived from studies of hyperfine structure of atomic lines and of alternating intensities in the rotational structure of bands of homonuclear diatomic molecules, whilst for the recognition of stable isotopes and determinations of their relative masses and abundances the isotope effect in diatomic band spectra, especially, has supplemented the mass spectrograph method to a large and important extent. One of these fruitful branches of modern spectroscopy is lucidly described in this monograph, largely with the aid of vector and energy-level diagrams and reference to actual examples. An account of electron-spin structure leads up to that of nuclear-spin structure. Derived values of the nuclear-spin quantum number I and magnetic-moment factor $g(I)$ are tabulated and discussed, and theories of nuclear spin are outlined. The reviewer ventures to express two hopes: (i) that nobody will use that "convenient unit" the millicentimetre^{-1} (p. 86)—what advantage has, say, "72 mcm^{-1}" over the familiar "0·072 cm^{-1}"? (ii) That a second edition will soon be called for, which will enable the author to correct a number of slips that have been found. It is usually stated that I is half-integer or zero according as the mass number is odd or even, except for ^{2}H and ^{14}N, which have I equal to 1. The author does well to emphasize that although this does describe the experimental results obtained up to the present, it should not be accepted as something in the nature of a *law*. It is true that lines of even-mass isotopes show no structure, and from this the spin is assumed to be zero in every case. But the nuclear masses for which the spin is known to be zero, from band-spectrum analysis, are not merely even but multiples of 4. So there is much to be said for the conclusion that I is half-integer or integer or zero according as the mass number is odd or an odd multiple of 2 or an even multiple of 2; to this there is no known exception.

W. J.